GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:14:11 ; Search time 2037.42 Seconds
(without alignments)
12546.865 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		781.4	41.3	799	10	BI770453	BI770453 603055418
2		712.4	37.6	730	10	BI772270	BI772270 603055851
3		682.6	36.0	991	10	BG179568	BG179568 602328182
c 4		659.6	34.8	721	9	AI745526	AI745526 wc34a10.x
c 5		617.4	32.6	640	10	BF589055	BF589055 nac29d04.
6		604.4	31.9	843	10	BG178499	BG178499 602328333
c 7		596.6	31.5	627	10	BF476482	BF476482 naa25e07.
8		590	31.2	933	10	BE563288	BE563288 601335410
9		584.8	30.9	945	10	BE871340	BE871340 601449457
c 10		576.4	30.4	625	10	BF446190	BF446190 7p33f08.x
11		573	30.3	590	10	BM006343	BM006343 603615035
12		549.8	29.0	899	10	BG328411	BG328411 602427749
c 13		549.6	29.0	606	9	AA662164	AA662164 ns66c07.s
c 14		525.8	27.8	652	9	AW130657	AW130657 xe64c07.x
15		504.4	26.6	514	9	AA662204	AA662204 ns66c07.r
16		467.6	24.7	809	10	BG967992	BG967992 602832361
17		465	24.6	601	9	AW971121	AW971121 EST383208
c 18		463.8	24.5	512	9	AI668916	AI668916 wb96d02.x
19		463.8	24.5	737	10	BI077619	BI077619 602871720
20		460.6	24.3	539	10	BE757615	BE757615 212104 MA
21		448.6	23.7	591	10	BG085247	BG085247 H3110C12-
c 22		431.8	22.8	458	9	AI991223	AI991223 wu41b03.x
c 23		417.4	22.0	430	9	AI979170	AI979170 wr71e06.x
c 24		396.2	20.9	418	9	AW298436	AW298436 UI-H-BW0-
25		381	20.1	943	10	BF676137	BF676137 602084129
c 26		372.2	19.7	392	9	AA639821	AA639821 nq84g11.s
c 27		352.6	18.6	361	9	AW513942	AW513942 xt83a04.x
c 28		349.8	18.5	353	10	BF003063	BF003063 7g53e04.x
c 29		346.8	18.3	381	9	AI859010	AI859010 wl45f04.x
30		345.6	18.2	386	9	AI685592	AI685592 tu20c07.x
c 31		345.6	18.2	548	9	AW973391	AW973391 EST385534
c 32		344	18.2	366	9	AI687669	AI687669 tp82g09.x
33		341.6	18.0	2064	10	BF579116	BF579116 602096050
c 34		323.6	17.1	366	9	AA662521	AA662521 nu99e07.s
35		316	16.7	355	9	AA513274	AA513274 ng94e10.s
c 36		306.8	16.2	384	9	AA548598	AA548598 nk54c12.s
37		302.6	16.0	859	10	BG916428	BG916428 602814031
38		297.4	15.7	628	10	BJ072997	BJ072997 BJ072997
39		289.6	15.3	369	9	AA228304	AA228304 nc38e07.r
c 40		278.6	14.7	784	10	BG072422	BG072422 H3110C12-
c 41		275.4	14.5	441	9	AI686414	AI686414 tu41g05.x
42		256.6	13.5	504	9	AA930371	AA930371 vo45a05.r
43		236.6	12.5	952	10	BF539428	BF539428 602049614
c 44		235	12.4	348	9	AI922279	AI922279 wn90c03.x
45		231	12.2	875	10	BG966737	BG966737 602834605

ALIGNMENTS

RESULT 1
BI770453
LOCUS BI770453 799 bp mRNA linear EST 25-SEP-2001
DEFINITION 603055418F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204692 5', mRNA sequence.
ACCESSION BI770453
VERSION BI770453.1 GI:15762031
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 799)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11513 row: d column: 05
High quality sequence stop: 799.
FEATURES Location/Qualifiers
source 1. .799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5204692"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size

```
                    range 1-3 kb. Library is normalized and enriched for
                    full-length clones and was  constructed by C. Gruber
                    (Invitrogen).  Research Genetics tracking  code 026. Note:
                    this is a NIH_MGC Library."
BASE COUNT      135 a    297 c    248 g    119 t
ORIGIN


  Query Match             41.3%;  Score 781.4;  DB 10;  Length 799;
  Best Local Similarity   98.6%;  Pred. No. 5.7e-135;
  Matches  788;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 180

Qy    181 gccaccagccacccctttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 240

Qy    241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    301 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    661 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 780

Qy    781 gcagtccatggtggtgggc 799
          ||||||||||||||||||
Db    781 GCAGTCCATGGTGGTGGGC 799


RESULT   2
BI772270
LOCUS       BI772270                730 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603055851F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5205132 5',
            mRNA sequence.
ACCESSION   BI772270
VERSION     BI772270.1  GI:15763848
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 730)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11514  row: f  column: 13
            High quality sequence stop: 729.
FEATURES             Location/Qualifiers
     source          1. .730
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5205132"
                     /clone_lib="NIH_MGC_122"
                     /lab_host="DH10B"
                     /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
                     anonymous pool of 24 week female lung, 16 week female
                     spleen, and 20-22 week male spleens.  Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon  cloning).  Average insert size 1.4 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was  constructed by C. Gruber
                     (Invitrogen).  Research Genetics tracking  code 026. Note:
                     this is a NIH_MGC Library."
BASE COUNT      124 a    265 c    237 g    104 t
ORIGIN


  Query Match             37.6%;  Score 712.4;  DB 10;  Length 730;
  Best Local Similarity   98.5%;  Pred. No. 3.3e-122;
  Matches  719;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     68 caagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccagcag 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTGCCAGCAG 60

Qy    128 gcagccctgggctgggggtaggggactccctacaggcacgcagccctgagaccgccacca 187
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     61 GCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACCTCAGAGG 120

Qy    188 gccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgccacca 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCACCCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCTGCCACCA 180

Qy    248 gccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgccagtggc 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGCCAGTGGC 240

Qy    308 ctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacagctccctg 367
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    241 CTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAGCTCCCTG 300

Qy    368 ggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgggcagcgcc 427
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGGCAGCGCC 360

Qy    428 agcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacacggtgtcg 487
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACACGGTGTCG 420

Qy    488 cggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggactggagt 547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGACTGGAGT 480

Qy    548 cccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcctactttgac 607
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    481 CCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTACTTTGAC 540

Qy    608 atgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcagtcgggag 667
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAGTCGGGAG 600

Qy    668 gagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagcgggcacc 727
          ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||| |
Db    601 GAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGCGGGCAGC 660

Qy    728 ctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggtgcagtcc 787
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGTGCAGTCC 720

Qy    788 atggtggtgg 797
          ||||||||||
Db    721 ATGGTGGTGG 730


RESULT   3
BG179568
LOCUS       BG179568                 991 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602328182F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429413 5',
            mRNA sequence.
ACCESSION   BG179568
VERSION     BG179568.1  GI:12686271
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 991)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10181  row: d  column: 22
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1. .991
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4429413"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      190 a    351 c    305 g    145 t
ORIGIN


  Query Match             36.0%;  Score 682.6;  DB 10;  Length 991;
  Best Local Similarity   95.6%;  Pred. No. 1.2e-116;
  Matches  735;  Conservative    0;  Mismatches   29;  Indels    5;  Gaps    3;

Qy    513 cgggggcagtgggtctcgagagacgggactggagtcccagtccacccgccacgcccgagc 572
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGGGGCAGTGGGTCTCGAGAGACGGGACTGGAGTCCCAGTCCACCCGCCACGCCCGAGC 60

Qy    573 agggcctgtccgccttcgacctctcctactttgacatgctgtaccctgaggacagcagct 632
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGGCCTGTCCGCCTTCTACCTCTCCTACTTTGACATGCTGTACCCTGAGGACAGCAGCT 120

Qy    633 gggcagccaaggcccctggggccagcagtcgggaggagccacctgaccagcctgagcagt 692
          |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    121 GGGCAGCCAAGGCCCCTGGGGCCAGCAGTCGGGAGGAGCCACCTGAGGAGCCTGAGCAGT 180

Qy    693 gcccggtcattgacagccaagccccagcgggcaccctggacttggtgcccggcgggctga 752
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    181 GCCCGGTCATTGACAGCCAAGCCCCAGCGGGCAGCCTGGACTTGGTGCCCGGCGGGCTGA 240

Qy    753 ccttggaggagcactcgctggagcaggtgcagtccatggtggtgggcgaagtgctcaagg 812
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTTGGAGGAGCACTCGCTGGAGCAGGTGCAGTCCATGGTGGTGGGCGAAGTGCTCAAGG 300

Qy    813 acatcgagacggcctgcaagctgctcaacatcaccgcagatcccatggactggagcccca 872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACATCGAGACGGCCTGCAAGCTGCTCAACATCACCGCAGATCCCATGGACTGGAGCCCCA 360

Qy    873 gcaatgtgcagaagtggctcctgtggacagagcaccaataccggctgccccccatgggca 932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCAATGTGCAGAAGTGGCTCCTGTGGACAGAGCACCAATACCGGCTGCCCCCCATGGGCA 420

Qy    933 aggccttccaggagctggcgggcaaggagctgtgcgccatgtcggaggagcagttccgcc 992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGGCCTTCCAGGAGCTGGCGGGCAAGGAGCTGTGCGCCATGTCGGAGGAGCAGTTCCGCC 480

Qy    993 agcgctcgcccctgggtggggatgtgctgcacgcccacctggacatctggaagtcagcgg 1052
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AGCGCTCGCCCCTGGGTGGGGATGTGCTGCACGCCCACCTGGACATCTGGAAGTCAGCGG 540

Qy   1053 cctggatgaaagagcggacttcacctggggcgattcactactgtgcctcgaccagtgagg 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTGGATGAAAGAGCGGACTTCACCTGGGGCGATTCACTACTGTGCCTCGACCAGTGAGG 600

Qy   1113 agagctggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctgtggc 1172
          |||||||||||||||||   |||||||||||||||||||||||||    |  |||||||
Db    601 AGAGCTGGACCGACAGCCGAGTGGACTCATCATGCTCCGGGCAGCC--ATCCACTGTGGC 658

Qy   1173 agttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtggctca 1232
          |||||||||||  |||||||||||||||||||||||||||||||||||||||||||  ||   |
Db    659 AGTTCCTCAAGCAGTTGCTACTCAAG-CCCACAGCTATGGCCGCTTCATAAG--TGGCTC 715

Qy   1233 acaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccg 1281
          || | ||||| ||| ||||| | ||||||||||||  ||| | | ||||
Db    716 ACCACGAGAACGGCCTCTTCCACATTGAGGACTCACGCCACGGGCCCCG 764


RESULT   4
AI745526/c
LOCUS       AI745526                 721 bp    mRNA    linear   EST 17-DEC-1999
DEFINITION  wc34a10.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2317050 3'
            similar to SW:ETS4_DROME P29775 DNA-BINDING PROTEIN D-ETS-4 ;, mRNA
            sequence.
ACCESSION   AI745526
VERSION     AI745526.1  GI:5113814
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 721)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
```

```
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 826   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 448.
FEATURES             Location/Qualifiers
     source          1. .721
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2317050"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      122 a    197 c    241 g    156 t      5 others
ORIGIN
```

Query Match 34.8%; Score 659.6; DB 9; Length 721;
Best Local Similarity 97.0%; Pred. No. 2e-112;
Matches 701; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

```
Qy   1172 cagttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtggctc 1231
          ||||||||||||   |||||||||  |||||  ||||||||  |||||||||  |||||||||
Db    721 CAGTTCCTCAAGAGTTTGCTACTC-AGCCCNNCAGCTATGNCCGCTTCATAAGGTGGCTC 663

Qy   1232 aacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggc 1291
          |  ||||||||||||||| ||||||||||||||||||||||||||||||| |||||||||
Db    662 ACAAAGGAGAAGGGCATC-TCAAAATTGAGGACTCAGCCCAGGTGGCCCGCCTGTGGGGC 604

Qy   1292 atccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccagtct 1351
          ||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db    603 ATCCGCAAGAACCGTCCCGCAATGAACTACGACAAGCTGAGCCGCTCCATCCGCCAGT-A 545

Qy   1352 tacaagaagggcatcatccggaagccagacatctcccagcgcctcgtctaccagttcgtg 1411
          |||||||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db    544 TACAAGAAGGGCATCATCCGNAAGCCAGACATCTCNCAGCGCCTCGTCTACCAGTTCGTG 485

Qy   1412 caccccatctgagtgcctggcccagggcctgaaacccgccctcaggggcctctctcctgc 1471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CACCCCATCTGAGTGCCTGGCCCAGGGCCTGAAACCCGCCCTCAGGGGCCTCTCTCCTGC 425

Qy   1472 ctgccctgcctcagccaggccctgagatgggggaaaacgggcagtgtgctctgctgctct 1531
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    424 CTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAAACGGGCAGTCTGCTCTGCTGCTCT 365

Qy   1532 gaccttccagagcccaaggtcagggaggggcaaccaactgccccaggggatatgggtcc 1591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 GACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCAACTGCCCCAGGGGGATATGGGTCC 305

Qy   1592 tctggggccttcgggaccatggggcaggggtgcttcctcctcaggcccagctgctccccct 1651
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    304 TCTGGGGCCTTCGGGACCCTGGGGCAGGGGTGCTTCCTCCTCAGGCCCAGCTGCTCCCCT 245

Qy   1652 ggaggacagagggagacagggctgctccccaacacctgcctctgaccccagcatttccag 1711
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 GGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACCTGCCTCTGACCCCAGCATTTCCAG 185

Qy   1712 agcagagcctacagaagggcagtgactcgacaaaggccacaggcagtccaggcctctctc 1771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 AGCAGAGCCTACAGAAGGGCAGTGACTCGACAAAGGCCACAGGCAGTCCAGGCCTCTCTC 125

Qy   1772 tgctccatccccctgcctcccattctgcaccacacctggcatggtgcagggagacatctg 1831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 TGCTCCATCCCCCTGCCTCCCATTCTGCACCACACCTGGCATGGTGCAGGGAGACATCTG 65

Qy   1832 caccctcagttgggcagccaggagtgccccgggaatggataataaagatactagagaa 1891
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     64 CACCCCTGAGTTGGGCAGCCAGGAGTGCCCCCGGGAATGGATAATAAAGATACTAAAGAA 5

Qy   1892 ctg 1894
          |||
Db      4 CTG 2
```

```
RESULT   5
BF589055/c
LOCUS       BF589055                640 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  nac29d04.x1 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:3394614 3' similar to TR:O95238 O95238 ETS TRANSCRIPTION
            FACTOR PDEF. ;, mRNA sequence.
ACCESSION   BF589055
VERSION     BF589055.1  GI:11681379
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lupski
             cDNA Library Preparation: Lupski Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 416.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3394614"
                     /clone_lib="Lupski_sciatic_nerve"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      103 a    182 c    218 g    136 t      1 others
ORIGIN
```

```
  Query Match            32.6%;  Score 617.4;  DB 10;  Length 640;
  Best Local Similarity  98.1%;  Pred. No. 1.3e-104;
  Matches  624;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1259 gaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaac 1318
          ||||||||||||  ||||||||| |||||||||||| ||||||||||||||||||||||
Db    640 GAGGACTCAGCCAGGGTGGCCCGACTGTGGGGCATCNGCAAGAACCGTCCCGCCATGAAC 581

Qy   1319 tgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagcca 1378
          | |||||||||||||||||||||||||||| |   ||||||||||||||||||||||||
Db    580 TACGACAAGCTGAGCCGCTCCATCCGCCAGTATACAAAGAAGGGCATCATCCGGAAGCCA 521

Qy   1379 gacatctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccaggg 1438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    520 GACATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGG 461

Qy   1439 cctgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgaga 1498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    460 CCTGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGA 401

Qy   1499 tgggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggag 1558
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    400 TGGGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAG 341

Qy   1559 gggcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcag 1618
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    340 GGGCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAG 281

Qy   1619 gggtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctc 1678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    280 GGGTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTC 221

Qy   1679 cccaacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgact 1738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 CCCAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACT 161

Qy   1739 cgacaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctg 1798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 CGACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTG 101

Qy   1799 caccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtg 1858
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    100 CACCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTG 41

Qy   1859 cccccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||||
Db     40 CCCCCGGGAATGGATAATAAAGATACTAGAGAACTG 5


RESULT   6
BG178499
LOCUS       BG178499                 843 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602328333F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429712 5',
            mRNA sequence.
ACCESSION   BG178499
VERSION     BG178499.1  GI:12685214
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10182  row: a  column: 09
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .843
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4429712"
                     /clone_lib="NIH_MGC_91"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      210 a    260 c    243 g    130 t
ORIGIN


  Query Match            31.9%;  Score 604.4;  DB 10;  Length 843;
  Best Local Similarity  94.9%;  Pred. No. 3.4e-102;
  Matches  647;  Conservative    0;  Mismatches   31;  Indels    4;  Gaps    2;

Qy   1213 ccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagccca 1272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGAGGACTCAGCCCA 60

Qy   1273 ggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgag 1332
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      61 GGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTACGACAAGCTGAG 120

Qy   1333 ccgctccatccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcg 1392
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     121 CCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGACATCTCCCAGCG 180

Qy   1393 cctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaaacccgccc 1452
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCCTGAAACCCGCCC 240

Qy   1453 tcaggggcctctctcctgcctgccctgcctcagccaggccctgagatgggggaaaacggg 1512
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAAACGGG 300

Qy   1513 cagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggggcaaccaactgc 1572
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCAACTGC 360

Qy   1573 cccagggggatatgggtcctctggggccttcgggaccatggggcaggggtgcttcctcct 1632
          |||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     361 CCCA-GGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGGGTGCTTCCTCCT 419

Qy   1633 caggcccagctgctcccctggaggacagagggagacagggctgctccccaacacctgcct 1692
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 CAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACCTGCCT 479

Qy   1693 ctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaaaggccaca 1752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 CTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAAAGGCCACA 539

Qy   1753 ggcagtccaggcctctctctgctccatcccccctgcctcccattctgcaccacacctggca 1812
          ||||||||||   ||||||   |    ||||  ||||||||||||||    | |
Db     540 GGCAGTCCAGG---CTCTCTCTGCTCATCCCCCTGCTCCCATTCTGCACACACCTGGCAT 596

Qy   1813 tggtgcagggagacatctgcacccctcagttgggcagccaggagtgcccccgggaatgga 1872
          |    |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     597 GGTGCAAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCCCCCGGGAATGGA 656

Qy   1873 taataaagatactagagaactg 1894
          | |||| ||||||||||||||||
Db     657 TCATAACGATACTAGAGAACTG 678
```

```
RESULT   7
BF476482/c
LOCUS       BF476482                627 bp    mRNA    linear   EST 05-DEC-2000
DEFINITION  naa25e07.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3256020 3'
            similar to TR:O95238 O95238 ETS TRANSCRIPTION FACTOR PDEF. ;, mRNA
            sequence.
ACCESSION   BF476482
VERSION     BF476482.1  GI:11547309
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 627)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 464.
FEATURES             Location/Qualifiers
     source          1. .627
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3256020"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      106 a    176 c    213 g    132 t
ORIGIN


  Query Match           31.5%;  Score 596.6;  DB 10;  Length 627;
  Best Local Similarity   97.0%;  Pred. No. 8.9e-101;
  Matches  608;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   1268 gcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaag 1327
           || |||||||||||||||||  |||| || | | | |||||||| |||||||| ||||||
Db    627 GCACAGGTGGCCCGGCTGTGGACCATCAGCTAAACCTGTCCCGCAATGAACTACGACAAA 568

Qy   1328 ctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccagacatctcc 1387
            |||||||||||||||||  || ||||||||||||||||||||||||||||||||||||
Db    567 GTGAGCCGCTCCATCCGCATGTATTACAAGAAGGGCATCATCCGGAAGCCAGACATCTCC 508

Qy   1388 cagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaaacc 1447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 CAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCCTGAAACC 448

Qy   1448 cgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatgggggaaa 1507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 CGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAA 388

Qy   1508 acgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggggcaacca 1567
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 ACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCA 328

Qy   1568 actgccccaggggatatgggtcctctggggccttcgggaccatggggcaggggtgcttc 1627
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    327 ACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGGGTGCTTC 268

Qy   1628 ctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccccaacacc 1687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 CTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACC 208

Qy   1688 tgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaaagg 1747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 TGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAAAGG 148

Qy   1748 ccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgcaccacacc 1807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 CCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCACCACACC 88

Qy   1808 tggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgccccccggga 1867
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     87 TGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCCCCCGGGA 28

Qy   1868 atggataataaagatactagagaactg 1894
           |||||||||||||||| ||||||||||
Db     27 ATGGATAATAAAGATAATAGAGAACTG 1


RESULT   8
BE563288
LOCUS       BE563288                933 bp    mRNA    linear   EST 15-AUG-2000
DEFINITION  601335410F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3689402 5',
            mRNA sequence.
ACCESSION   BE563288
VERSION     BE563288.1  GI:9807100
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM384  row: c  column: 03
            High quality sequence stop: 678.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3689402"
                     /clone_lib="NIH_MGC_39"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Library constructed
                     by Ling Hong in the laboratory of Gerald M. Rubin
                     (University of California,  Berkeley) using ZAP-cDNA
                     synthesis kit (Stratagene) and Superscript II RT (Life
                     Technologies)."
BASE COUNT      148 a    316 c    319 g    150 t
ORIGIN
```

```
  Query Match            31.2%;  Score 590;  DB 10;  Length 933;
  Best Local Similarity   95.0%;  Pred. No. 1.6e-99;
  Matches  663;  Conservative    0;  Mismatches   30;  Indels    5;  Gaps    5;

Qy      21 attcctgcactctgccgtgtccacactgccccacagacccagtcctccaagcctgctgcc 80
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1 ATTCCTGCACTCTGCCGTGT-CACACTGCCCCACAGACCCAGTCCTCCAAGCCTGCTGCC 59

Qy      81 agctccctgcaagcccctcaggttgggccttgccacggtgccagcaggcagccctgggct 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60 AGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTGCCAGCAGGCAGCCCTGGGCT 119

Qy     141 gggggtaggggactccctacaggcacgcagccctgagaccgccaccagccacccccttgag 200
           |||||||||||||||||||||||||||||||||||||| |     ||||||||||||
Db     120 GGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACCTCAGAGGGCCACCCCTTGAG 179

Qy     201 ggtggccaggcccccagtggccaacctgagtgctgcctctgccaccagccctgctggccc 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 GGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCTGCCACCAGCCCTGCTGGCCC 239

Qy     261 ctggttccgctggccccccagatgcctggctgagacacgccagtggcctcagctgcccac 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 CTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGCCAGTGGCCTCAGCTGCCCAC 299

Qy     321 acctcttcccggcccctggagttggcactgcagcagacagctccctgggcaccaggcagc 380
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db     300 ACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAGCTCCCTGGGCACCAGGCAGC 359

Qy     381 taacagacacagccgccagcccaaacagcagcggcatgggcagcgccagcccgggtctga 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGGCAGCGCCAGCCCGGGTCTGA 419

Qy     441 gcagcgtatcccccagccacctcctgctgccccccgacacggtgtcgcggacaggcttgg 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 GCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACACGGTGTCGCGGACAGGCTTGG 479

Qy     501 agaaggcggcagc-gggggcagtgggtctcgagagacgggactggagtcccagtccaccc 559
           |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     480 AGAAGGCGGCAGCGGGGGGCAGTGGGTCTCGAGAGACGGGACTGGAGTCCCAGTCCACCC 539

Qy     560 gccacgcccgagcagggcctgtccgccttcgacctctcctactttgacatgctgtaccct 619
           |||||||||||||||| ||||||| |    |    |    |||||||||||||||||||
Db     540 GCCACGCCCGAGCAGGG-CTGTCCGCTTCTACCTCTCCTACTTTTGACATGCTGTACCCT 598

Qy     620 gaggacagcagctgggcagccaaggcccctggggccagcagtcgggaggagccacctgac 679
           ||||||||||||  ||    |||||||||||||| ||||||||||||||||||||||||
Db     599 GAGGACAGCAGCCTGGGCAGCAAGGCCCCTGGGG-CAGCAGTCGGGAGGAGCCACCTGAG 657

Qy     680 cagcc-tgagcagtgcccggtcattgacagccaagccc 716
           |||| |||||||||||||||||||| | |||||||
Db     658 GAGCCTTGAGCAGTGCCCGGTCATTGACTGACAAGCCC 695


RESULT   9
BE871340
LOCUS       BE871340                945 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601449457F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3853334 5',
            mRNA sequence.
ACCESSION   BE871340
VERSION     BE871340.1  GI:10320116
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 945)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9577  row: a  column: 15
            High quality sequence stop: 656.
FEATURES             Location/Qualifiers
     source          1. .945
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3853334"
                     /clone_lib="NIH_MGC_65"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies.  "
BASE COUNT      170 a    363 c    300 g    112 t
ORIGIN


  Query Match            30.9%;  Score 584.8;  DB 10;  Length 945;
  Best Local Similarity   90.6%;  Pred. No. 1.5e-98;
  Matches  703;  Conservative    0;  Mismatches   62;  Indels   11;  Gaps    7;

Qy      51 ccacagacccagtcctccaagcctgctgccagctccctgcaagcccctcaggttgggcct 110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCACAGACCCAGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCT 60

Qy     111 tgccacggtgccagcaggcagccctgggctgggggtaggggactccctacaggcacgcag 170
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      61 TGCCACGGTGCCAGCAGGCAG-CCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAG 119

Qy     171 ccctgagaccgccaccagccacccccttgagggtggccaggcccccagtggccaacctgag 230
           |||||||||| |     ||||||||||||||||||||||||||||||||||||||||||||
Db     120 CCCTGAGACCTCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAG 179

Qy     231 tgctgcctctgccaccagccctgctggcccctggttccgctggccccccagatgcctggc 290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 TGCTGCCTCTGCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGC 239

Qy     291 tgagacacgccagtggcctcagctgcccacacctcttcccggcccctggagttggcactg 350
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     240 TGAGACACGCCAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTG 299

Qy     351 cagcagacagctccctgggcaccaggcagctaacagacacagccgccagcccaaacagca 410
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 CAGCAGACAGCTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCA 359

Qy     411 gcggcatgggcagcgccagcccgggtctgagcagcgtatccccccagccacctcctgctgc 470
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 GCGGCATGGGCAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGC 419

Qy     471 cccccgacacggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcg 530
           |||||||||||| |||||||||||||||||||||||| |||||||||| |||||||||
Db     420 CCCCCGACACGG-GTCGCGGACAGGCTTGGAGAAGGCGG-AGCGGGGGCAGCGGGTCTCG 477

Qy     531 agagacgggactggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcg 590
           |||||||||||||||||||||||||||||||||||||   |     | |    |  ||||
Db     478 AGAGACGGGACTGGAGTCCCAGTCCACCCGCCACGCC---GAGCAGGCCTGTCCGCTTCT 534

Qy     591 acctctcctactttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctg 650
           |||||||||||     |||||||| ||||||||||||||||||||||| | | |||  |
Db     535 ACCTCTCCTACTCGACCATGCTGTCACCTGAGGACAGCAGCTGGGCAGCAAGGCCCCTGG 594

Qy     651 gggccagcagtcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagcc 710
           || |  ||||||||||||||||||||  |||||||||| |||||||||    |
Db     595 GGCCAGCCAGTCGGGAGGAGCCACCTGAGGAGCCTGAGCAGGGCCCGGTCAT---GACAG 651
```

```
Qy    711 aagccccagcgggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgc 770
          |  |||||| |     ||| ||||||||||||| |||||||||||||||||||||||||
Db    652 CAAGCCCAGCTGCGCAGCTGCACTTGGTGCCCGGC-GGCTGACCTTGGAGGAGCACTCGC 710

Qy    771 tggagcaggt-gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggc 825
          |||||||||| || ||||||||||||||||||| |||| || |     ||||||| |
Db    711 TGGAGCAGGTGGCCGTCCATGGTGGTGGGCGAGGTGCCCACGACTTCCGAGACGCC 766


RESULT 10
BF446190/c
LOCUS       BF446190                625 bp    mRNA    linear   EST 01-DEC-2000
DEFINITION  7p33f08.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3647751 3'
            similar to TR:O95238 O95238 ETS TRANSCRIPTION FACTOR PDEF. ;, mRNA
            sequence.
ACCESSION   BF446190
VERSION     BF446190.1  GI:11511328
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 460.
FEATURES             Location/Qualifiers
     source          1. .625
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3647751"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      113 a    174 c    202 g    135 t      1 others
ORIGIN


  Query Match           30.4%;  Score 576.4;  DB 10;  Length 625;
  Best Local Similarity   97.1%;  Pred. No. 4.9e-97;
  Matches  597;  Conservative    0;  Mismatches   17;  Indels    1;  Gaps    1;

Qy   1280 cggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctcc 1339
          |||||||||||||||||||| |||||||||||||||||  | ||||||||||||||||
Db    625 CGGCTGTGGGGCATCCGCAAGANCCGTCCCGCCATGAACTA-GCCAAGCTGAGCCGCTCC 567

Qy   1340 atccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgtc 1399
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    566 ATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGACATCTCCCAGCGCCTCGTC 507

Qy   1400 taccagttcgtgcaccccatctgagtgcctggcccagggcctgaaacccgccctcagggg 1459
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    506 TACCAGTTCGTGCACCCCATTTGAGTGCCTGGCCCAGGGCCTGAAACCCGCCCTCAGGGG 447

Qy   1460 cctctctcctgcctgccctgcctcagccaggccctgagatgggggaaaacgggcagtgtg 1519
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    446 CCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAAACGGGCAGTTTG 387

Qy   1520 ctctgctgctctgaccttccagagcccaaggtcagggaggggcaaccaactgccccaggg 1579
          || ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 CTTTGCTGCTTTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCAACTGCCCCAGGG 327

Qy   1580 ggatatgggtcctctggggccttcgggaccatggggcaggggtgcttcctcctcaggccc 1639
          |||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||
Db    326 GGATATGGGTCCTTTGGGGCCTTCGGGACCCTGGGGCAGGGGTGCTTCCTCCTCAGGCCC 267

Qy   1640 agctgctcccctggaggacagagggagacagggctgctccccaacacctgcctctgaccc 1699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    266 AGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACCTGCCTTTGACCC 207

Qy   1700 cagcatttccagagcagagcctacagaagggcagtgactcgacaaaggccacaggcagtc 1759
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 CAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAAAGGCCACAGGCAGTC 147

Qy   1760 caggcctctctctgctccatccccctgcctcccattctgcaccacacctggcatggtgca 1819
          ||||||| | | ||||||||||||||||||||||| ||||||||||||||||||||||||
Db    146 CAGGCCTTTTTTTGCTCCATCCCCCTGCCTCCCATTTTGCACCACACCTGGCATGGTGCA 87

Qy   1820 gggagacatctgcacccctcagttgggcagccaggagtgccccccgggaatggataataaa 1879
          ||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     86 GGGAGACATTTGCACCCCTGAGTTGGGCAGCCAGGAGTGCCCCCGGGAATGGATAATAAA 27

Qy   1880 gatactagagaactg 1894
          |||||||||||||||
Db     26 GATACTAGAGAACTG 12


RESULT 11
BM006343
LOCUS       BM006343                590 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603615035F1 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5420913 5',
            mRNA sequence.
ACCESSION   BM006343
VERSION     BM006343.1  GI:16520697
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1876  row: e  column: 10
            High quality sequence stop: 578.
FEATURES             Location/Qualifiers
     source          1. .590
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5420913"
                     /clone_lib="NIH_MGC_110"
                     /tissue_type="ductal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
```

```
                Site_2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."
BASE COUNT      101 a    225 c    180 g     84 t
ORIGIN


  Query Match             30.3%;  Score 573;  DB 10;  Length 590;
  Best Local Similarity   98.3%;  Pred. No. 2.1e-96;
  Matches  579;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      60 cagtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggt 119
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 CAGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGT 61

Qy     120 gccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagac 179
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 GCCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGAC 121

Qy     180 cgccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctc 239
           | |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 CTCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTC 181

Qy     240 tgccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacg 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 TGCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACG 241

Qy     300 ccagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagaca 359
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     242 CCAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACA 301

Qy     360 gctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgg 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 GCTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGG 361

Qy     420 gcagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccccgaca 479
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     362 GCAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGACCCCCGACA 421

Qy     480 cggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggg 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 CGGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGG 481

Qy     540 actggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcct 599
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     482 ACTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCT 541

Qy     600 actttgacatgctgtaccctgaggacagcagctgggcagccaaggcccc 648
           |||||||||||||||||||||||||||||||||||| |||||||||
Db     542 ACTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGGCAAGGCCCC 590


RESULT  12
BG328411
LOCUS       BG328411                 899 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602427749F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4546860 5',
            mRNA sequence.
ACCESSION   BG328411
VERSION     BG328411.1  GI:13134849
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 899)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1234  row: b  column: 13
            High quality sequence stop: 602.
FEATURES             Location/Qualifiers
     source          1. .899
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4546860"
                     /clone_lib="NIH_MGC_15"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      163 a    346 c    277 g    113 t
ORIGIN


  Query Match             29.0%;  Score 549.8;  DB 10;  Length 899;
  Best Local Similarity   86.9%;  Pred. No. 4.5e-92;
  Matches  734;  Conservative    0;  Mismatches   92;  Indels   19;  Gaps   11;

Qy       5 gacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtc 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 GACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCCAGTC 61

Qy      65 ctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccag 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 CTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTGCCAG 121

Qy     125 caggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagaccgcca 184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     122 CAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACCTCAG 181

Qy     185 ccagccacccctttgagggtggccaggcccccagtggccaacctgagtgctgcctctgcca 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 AGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCTGCCA 241

Qy     245 ccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgccagt 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 CCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGCCAGT 301

Qy     305 ggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacagctcc 364
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     302 GGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAGCTCC 361

Qy     365 ctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgggcagc 424
           |||||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db     362 CTGGGCACCAGGCAGCTAACAGACACAGTCG-CAGCCCAAACAGCAGCGGCATGGGCAGC 420

Qy     425 gccagcccggg-tctgagcagcgtatcccccagccacctcctgctgccccccgacacggt 483
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GCCAGCCCGGGCTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACACGG- 479

Qy     484 gtcgcggacaggcttggagaa-ggcggcagcgggggcagtgggtctcgagagacgggact 542
           |||||||||||||||||||| |||||||||||||||| |||||||||||||||||||||
Db     480 GTCGCGGACAGGCTTGGAGAACGGCGGCAGCGGGGGCAGCGGGTCTCGAGAGACGGGACT 539

Qy     543 ggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcctact 602
           ||||||||||||||||||     |   |    | |   |||||||| ||||||||||
Db     540 GGAGTCCCAGTCCACCCGC----ACGCCGAGCAGGCCTGCCGCCTTCTACCTCTCCTAC- 594
```

```
Qy    603 ttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcagtc 662
          ||||||||||||||||||||||||  |  |||||   | ||||||||  | | |
Db    595 TTGACATGCTGTACCCTGAGGACCGACGCTGGCACGGCAAGGCCCTGGCGCCAGCAGCCG 654

Qy    663 gggaggagccacctga-ccagcctgagcagtgcccggtcattgacagccaagccccagcg 721
          || |||||||||||||| | ||||||||||||||||  |  ||||||||||  |||| |
Db    655 GGAAGGAGCCACCTGAGCGAGCCTGAGCAGTGCCCG--CCATGACAGCCAAG--CCAGGG 710

Qy    722 ggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggtg 781
          ||||  |||||||||||||||     ||||   | ||||||||||  ||||||||||
Db    711 GGCAGCCTGGACTTGGTGCC---GCGGCTAGACCTGGAGGAGCA--CGCTGGAGCACCGT 765

Qy    782 cagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaac 841
          |||||  |  |||||||   |  |  |    ||   |  |   ||  |    |||
Db    766 GCAGCCATGTCGCGGGGCGAACGCCAGACTGCAGACGCCTCCACTGCCACTCCCGCGAAC 825

Qy    842 atcac 846
          || ||
Db    826 ATGAC 830
```

```
RESULT 13
AA662164/c
LOCUS       AA662164                606 bp    mRNA    linear   EST 03-DEC-1997
DEFINITION  ns66c07.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1188588 3'
            similar to SW:ETS4_DROME P29775 DNA-BINDING PROTEIN D-ETS-4 ;, mRNA
            sequence.
ACCESSION   AA662164
VERSION     AA662164.1  GI:2616255
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 606)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 734   Std Error: 0.00
            Seq primer: -40ml3 fwd. ET from Amersham
            High quality sequence stop: 393.
FEATURES             Location/Qualifiers
     source          1. .606
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1188588"
                     /clone_lib="NCI_CGAP_Pr22"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I  and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      104 a    172 c    204 g    126 t
ORIGIN
```

```
  Query Match           29.0%;  Score 549.6;  DB 9;  Length 606;
  Best Local Similarity  97.2%;  Pred. No. 4.5e-92;
  Matches  591;  Conservative    0;  Mismatches   14;  Indels    3;  Gaps    3;

Qy   1288 gggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgcca 1347
          ||| |||||||||| |||||||||||||||| |||||||||||| |||||||| ||||
Db    606 GGGGATCCGCAAGAA-CGTCCCGCCATGAACTACGACAAGCTGAGCGGCTCCATCGGCCA 548

Qy   1348 gtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgtctaccagtt 1407
          || ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    547 GTATTACAAGAAGGGCATCATCCGGAAGCCAGACATTTCCCAGCGCCTCGTCTACCAGTT 488

Qy   1408 cgtgcaccccatctgagtgcctggcccagggcctgaaacccgccctcaggggcctctctc 1467
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    487 CGTGCACCCCATCTGAGTGCCTGGCCCAGGGCCTGAAACCCGGCCTCAGGGGCCTCTCTC 428

Qy   1468 ctgcctgccctgcctcagccaggccctgagatgggggaaaacgggcagtgtgctctgctg 1527
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    427 CTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAAACGGGCAGTCTGCTCTGCTG 368

Qy   1528 ctctgaccttccagagcccaaggtcagggaggggcaaccaactgccccaggggatatgg 1587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 CTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCAACTGCCCCAGGGGGATATGG 308

Qy   1588 gtcctctgggggccttcgggaccatggggcaggggtgcttcctcctcaggcccagctgctc 1647
          |||||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    307 GTCCTCTGGGG-CTTCGGGACCCTGGGGCAGGGGTGCTTCCTCCTCAGGCCCAGCTGCTC 249

Qy   1648 ccctggaggacagagggagacagggctgctccccaacacctgcctctgaccccagcattt 1707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 CCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACCTGCCTCTGACCCCAGCATTT 189

Qy   1708 ccagagcagagcctacagaagggcagtgactcgacaaaggccacaggcagtccaggcctc 1767
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    188 CCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAAGAGCCACAGGCAGTCCAGGCCTC 129

Qy   1768 tctctgctccatccccctgcctcccattctgcaccacacctggcatggtgcagggagaca 1827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 TCTCTGCTCCATCCCCCTGCCTCCCATTCTGCACCACACCTGGCATGGTGCAGGGAGACA 69

Qy   1828 tctgcacccctcagttgggcagccaggagtgcccccg-ggaatggataataaagatacta 1886
          |||||||||| ||||| ||||||| |||||||||| ||||||||||||||||||||||
Db     68 TCTGCACCCCTGAGTTGTGCAGCCATGAGTGCCCCCGTGGAATGGATAATAAAGATACTA 9

Qy   1887 gagaactg 1894
          ||||||||
Db      8 GAGAACTG 1
```

```
RESULT 14
AW130657/c
LOCUS       AW130657                652 bp    mRNA    linear   EST 27-OCT-1999
DEFINITION  xe64c07.x1 NCI_CGAP_Ut3 Homo sapiens cDNA clone IMAGE:2612652 3'
            similar to TR:O95238 O95238 ETS TRANSCRIPTION FACTOR PDEF. ;, mRNA
            sequence.
ACCESSION   AW130657
VERSION     AW130657.1  GI:6132353
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 652)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
```

```
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Seq primer: -40UP from Gibco
           High quality sequence stop: 423.
FEATURES             Location/Qualifiers
     source          1. .652
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2612652"
                     /clone_lib="NCI_CGAP_Ut3"
                     /tissue_type="poorly-differentiated endometrial
                     adenocarcinoma, 2 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.45 kb. Life Technologies catalog #:
                     11541-018"
BASE COUNT      116 a    182 c    210 g    142 t      2 others
ORIGIN


  Query Match             27.8%;  Score 525.8;  DB 9;  Length 652;
  Best Local Similarity   94.9%;  Pred. No. 1.2e-87;
  Matches  564;  Conservative    0;  Mismatches   28;  Indels    2;  Gaps    2;

Qy   1302 accgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagg 1361
           || |||||||||   | | ||||||| | | ||| |||||||||||||| |   | | ||
Db    600 ACGGTCCCGCCAATGAATACGACAAGTTAACCCG-TCCATCCGCCAGTATACAAGAAGGG 542

Qy   1362 gcatcatccggaagccagacatctcccagcgcctcgtctaccagttcgtgcaccccatct 1421
           |||||||| | |  ||||||||||||||||||| ||||||||||||||||||||||||
Db    541 GCATCATCGGAAGCCCAGACATCTCCCAGCGCTTCGTCTACCAGTTCGTGCACCCCATCT 482

Qy   1422 gagtgcctggcccagggcctgaaacccgccctcaggggcctctctcctgcctgccctgcc 1481
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    481 GAGTGCCTGGCCCAGGGCTTGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCC 422

Qy   1482 tcagccaggccctgagatgggggaaaacgggcagtgtgctctgctgctctgaccttccag 1541
           |||||||||||||||||||||||| ||||||||| || ||||||||||||||||||||
Db    421 TCAGCCAGGCCCTGAGATGGGGGAAANACGGGCAGTCTGTTCTGCTGCTCTGACCTTCCAG 362

Qy   1542 agcccaaggtcagggaggggcaaccaactgccccaggggatatgggtcctctggggcct 1601
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGCCCAAGGTCAGGGAGGGGCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCT 302

Qy   1602 tcgggaccatggggcaggggtgcttcctcctcaggcccagctgctcccctggaggacaga 1661
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCGGGACCCTGGGGCAGGGGTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGA 242

Qy   1662 gggagacagggctgctccccaacacctgcctctgaccccagca-tttccagagcagagcc 1720
           |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    241 GGGAGACAGGGCTGCTCCCCAACACCTGCCTCTGACCCCAGCATTTTCCAGAGCAGAGCC 182

Qy   1721 tacagaagggcagtgactcgacaaaggccacaggcagtccaggcctctctctgctccatc 1780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TACAGAAGGGCAGTGACTCGACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATC 122

Qy   1781 cccctgcctcccattctgcaccacacctggcatggtgcagggagacatctgcacccctca 1840
           |||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |
Db    121 CCCCTGCCTCCCATTTTGCACCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGA 62

Qy   1841 gttgggcagccaggagtgccccccgggaatggataataaagatactagagaactg 1894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTTGGGCAGCCAGGAGTGCCCCCGGGAATGGATAATAAAGATACTAGAGAACTG 8


RESULT  15
AA662204
LOCUS       AA662204                 514 bp    mRNA    linear   EST 03-DEC-1997
DEFINITION  ns66c07.r1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1188588 5'
            similar to SW:ETS4_DROME P29775 DNA-BINDING PROTEIN D-ETS-4 ;, mRNA
            sequence.
ACCESSION   AA662204
VERSION     AA662204.1  GI:2616295
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 514)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Insert Length: 734    Std Error: 0.00
           Seq primer: -28ml3 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1188588"
                     /clone_lib="NCI_CGAP_Pr22"
                     /sex="male"
                     /tissue_type="normal prostate"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; 1st strand cDNA was prepared
                     from normal prostate bulk tissue, and was then primed with
                     a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I  and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized, and was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      108 a    173 c    151 g     82 t
ORIGIN


  Query Match             26.6%;  Score 504.4;  DB 9;  Length 514;
  Best Local Similarity   98.8%;  Pred. No. 1e-83;
  Matches  508;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy   1240 gaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaa 1299
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGAGGGCATCTTCAAAATTGAGGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAA 60

Qy   1300 gaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaa 1359
           |||||||||||||||||| |||||||||||||||||||||||||||| |||||||||
Db     61 GAACCGTCCCGCCATGAACTACGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAA 120

Qy   1360 gggcatcatccggaagccagacatctcccagcgcctcgtctaccagttcgtgcaccccat 1419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGGCATCATCCGGAAGCCAGACATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCAT 180

Qy   1420 ctgagtgcctggcccagggcctgaaacccgccctcaggggcctctctcctgcctgccctg 1479
           |||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||
Db    181 CTGAGTGCCTGGACCAGGGCATGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTG 240

Qy   1480 cctcagccaggccctgagatgggggaaaacgggcagtgtgctctgctgctctgaccttcc 1539
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    241 CCTCAGCCAGGCCCTGAGATGGGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCC 300

Qy   1540 agagcccaaggtcagggaggggcaaccaactgccccaggggatatgggtcctctggggc 1599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     301 AGAGCCCAAGGTCAGGGAGGGGCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGC 360

Qy    1600 cttcgggaccatggggcaggggtgcttcctcctcaggcccagctgctcccctggaggaca 1659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CTTCGGGACCATGGGGCAGGGGTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACA 420

Qy    1660 gagggagacagggctgctccccaacacctgcctctgaccccagcatttccagagcagagc 1719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAGGGAGACAGGGCTGCTCCCCAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGC 480

Qy    1720 ctacagaagggcagtgactcgacaaaggccacag 1753
           |||||||||||||||||||||||||||||||||
Db     481 CTACAGAAGGGCAGTGACTCGACAAAGGCCACAG 514
```

Search completed: August 6, 2002, 18:52:17
Job time: 2286 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:17:56 ; Search time 377.48 Seconds
(without alignments)
10918.304 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1362792 seqs, 1088025756 residues

Total number of hits satisfying chosen parameters: 2725584

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:*
8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1870	98.7	1894	1	PCT-US02-18947-1406	Sequence 1406, Ap
	2	1870	98.7	1894	7	US-10-172-118-1406	Sequence 1406, Ap
	3	1864.4	98.4	1942	7	US-10-157-031-83	Sequence 83, Appl
	4	1019.8	53.8	1087	7	US-10-102-806-215	Sequence 215, App
c	5	790.6	41.7	1188	5	US-09-971-429B-54	Sequence 54, Appl
	6	80.4	4.2	2286	5	US-09-053-375B-626	Sequence 626, App
c	7	75	4.0	499	1	PCT-US01-43704-2290	Sequence 2290, Ap
	8	75	4.0	1907	1	PCT-US02-07826-74	Sequence 74, Appl
	9	75	4.0	1907	7	US-10-097-340-74	Sequence 74, Appl
	10	75	4.0	2457	5	US-09-919-002-13047	Sequence 13047, A
	11	75	4.0	2642	7	US-10-099-926-1680	Sequence 1680, Ap
	12	75	4.0	3045	5	US-09-654-788A-107	Sequence 107, App
	13	74.8	3.9	2064	6	US-10-167-485-1	Sequence 1, Appli
	14	74.8	3.9	2333	7	US-10-099-926-1788	Sequence 1788, Ap
	15	74	3.9	502	1	PCT-US02-12378-282	Sequence 282, App
	16	74	3.9	502	7	US-10-124-805-282	Sequence 282, App
	17	74	3.9	3126	5	US-09-442-384B-475	Sequence 475, App
	18	73.2	3.9	4071	1	PCT-US02-18947-1042	Sequence 1042, Ap
	19	73.2	3.9	4071	7	US-10-007-926A-300	Sequence 300, App
	20	73.2	3.9	4071	7	US-10-172-118-1042	Sequence 1042, Ap
	21	72.2	3.8	855	5	US-09-053-375B-12	Sequence 12, Appl
	22	71	3.7	2957	1	PCT-US02-18947-710	Sequence 710, App
	23	71	3.7	2957	7	US-10-007-926A-295	Sequence 295, App
	24	71	3.7	2957	7	US-10-172-118-710	Sequence 710, App
	25	70.4	3.7	317	7	US-10-007-926A-299	Sequence 299, App
	26	70.2	3.7	2410	5	US-09-053-375B-920	Sequence 920, App
	27	69.4	3.7	414	7	US-10-007-926A-369	Sequence 369, App
	28	69.4	3.7	1932	5	US-09-053-375B-327	Sequence 327, App
	29	69.4	3.7	1932	5	US-09-442-384B-530	Sequence 530, App
	30	69	3.6	2658	5	US-09-053-375B-634	Sequence 634, App
	31	66.2	3.5	1729	5	US-09-053-375B-627	Sequence 627, App
	32	65.2	3.4	251	5	US-09-930-213-353	Sequence 353, App
	33	65.2	3.4	1547	1	PCT-US02-18947-180	Sequence 180, App
	34	65.2	3.4	1547	7	US-10-172-118-180	Sequence 180, App
	35	65.2	3.4	1551	1	PCT-US02-18947-1610	Sequence 1610, Ap
	36	65.2	3.4	1551	7	US-10-172-118-1610	Sequence 1610, Ap
	37	63	3.3	473	5	US-09-918-995-27904	Sequence 27904, A
	38	63	3.3	2188	7	US-10-099-926-1716	Sequence 1716, Ap
	39	63	3.3	2268	7	US-10-099-926-1693	Sequence 1693, Ap
	40	63	3.3	2269	5	US-09-954-531-955	Sequence 955, App
	41	63	3.3	3692	7	US-10-106-698-824	Sequence 824, App
c	42	63	3.3	3777	5	US-09-919-002-3576	Sequence 3576, Ap
	43	62.6	3.3	518	8	US-60-377-240-2401	Sequence 2401, Ap
	44	60.4	3.2	659	7	US-10-027-632-23683	Sequence 23683, A
	45	60.4	3.2	692	8	US-60-377-240-3339	Sequence 3339, Ap

ALIGNMENTS

RESULT 1
PCT-US02-18947-1406
; Sequence 1406, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1406
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_012391
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-1406

Query Match 98.7%; Score 1870; DB 1; Length 1894;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
             |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
```

```
Db    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

RESULT 2
US-10-172-118-1406
; Sequence 1406, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1406
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_012391
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1406

Query Match 98.7%; Score 1870; DB 7; Length 1894;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
           |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tcagagggccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggccccctggttccgctggcccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 gccaccagccctgctggccccctggttccgctggcccccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccccgacac 480

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db     661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy    1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800

Qy    1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy    1861 cccgggaatggataataaagatactagagaactg 1894
           |||||||||||||||||||||||||||||||||||
Db    1861 cccgggaatggataataaagatactagagaactg 1894
```

RESULT 3
US-10-157-031-83
; Sequence 83, Application US/10157031
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequenc
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-157-031-83

Query Match 98.4%; Score 1864.4; DB 7; Length 1942;

```
Best Local Similarity   99.2%;  Pred. No. 0;
Matches 1874;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       5 gacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtc 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 gacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtc 69

Qy      65 ctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccag 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 ctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccag 129

Qy     125 caggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagaccgcca 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130 caggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacctcag 189

Qy     185 ccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgcca 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 agggccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgcca 249

Qy     245 ccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgccagt 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 ccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgccagt 309

Qy     305 ggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacagctcc 364
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     310 ggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacagctcc 369

Qy     365 ctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgggcagc 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 ctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgggcagc 429

Qy     425 gccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacacggtg 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 gccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacacggtg 489

Qy     485 tcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggactgg 544
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 tcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggactgg 549

Qy     545 agtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcctacttt 604
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     550 agtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctcctacttt 609

Qy     605 gacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcagtcgg 664
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 gacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcagtcgg 669

Qy     665 gaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagcgggc 724
           ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db     670 gaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagcgggc 729

Qy     725 accctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggtgcag 784
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     730 agcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggtgcag 789

Qy     785 tccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaacatc 844
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     790 tccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaacatc 849

Qy     845 accgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggacagag 904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     850 accgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggacagag 909

Qy     905 caccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaaggagctg 964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     910 caccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaaggagctg 969

Qy     965 tgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgctgcac 1024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     970 tgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgctgcac 1029

Qy    1025 gcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctggggcg 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1030 gcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctggggcg 1089

Qy    1085 attcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactcatca 1144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1090 attcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactcatca 1149

Qy    1145 tgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccac 1204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1150 tgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccac 1209

Qy    1205 agctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggac 1264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1210 agctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggac 1269

Qy    1265 tcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgac 1324
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    1270 tcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgac 1329

Qy    1325 aagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccagacatc 1384
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    1330 aagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccagacatc 1389

Qy    1385 tcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaa 1444
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1390 tcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaa 1449

Qy    1445 acccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatggggg 1504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1450 acccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatggggg 1509

Qy    1505 aaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggggcaa 1564
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1510 aaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggggcaa 1569

Qy    1565 ccaactgccccagggggatatgggtcctctggggccttcgggaccatggggcaggggtgc 1624
           ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1570 ccaactgccccagggggatatgggtcctctggggccttcgggaccctggggcaggggtgc 1629

Qy    1625 ttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccccaac 1684
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1630 ttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccccaac 1689

Qy    1685 acctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaa 1744
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 acctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaa 1749

Qy    1745 aggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgcaccac 1804
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1750 aggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgcaccac 1809

Qy    1805 acctggcatggtgcagggagacatctgcaccccctcagttgggcagccaggagtgccccg 1864
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    1810 acctggcatggtgcagggagacatctgcaccccctgagttgggcagccaggagtgccccg 1869

Qy    1865 ggaatggataataaagatactagagaactg 1894
           ||||||||||||||||||||||||||||||
Db    1870 ggaatggataataaagatactagagaactg 1899
```

RESULT 4
US-10-102-806-215
; Sequence 215, Application US/10102806
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

```
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 215
;   LENGTH: 1087
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1075)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-215


  Query Match             53.8%;  Score 1019.8;  DB 7;  Length 1087;
  Best Local Similarity   97.0%;  Pred. No. 7.8e-184;
  Matches 1055;  Conservative    4;  Mismatches   27;  Indels    2;  Gaps    2;

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 61

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 121

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db     122 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 181

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 241

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 301

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 361

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 421

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 481

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 541

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 601

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 661

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 721

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db     722 ggactcagcccaggtggcccggctgtrgggcatccgcaagaaccgtcccgccatgaacta 781

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           ||||||||||||||||||||||||:|||| ||||||||||||||||||||||||||||||
Db     782 cgacaagctgagccgctccatccgscagtattacaagaagggcatcatccggaagccaga 841

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           |||||:||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 catctyccagcgsctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 901

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 961

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
           ||||||||| |||||| |||||||||||||||||||| |||||||||||||| ||||||
Db     962 ggggaaaac-ggcagtctgctctgctgctctgacctt-cagagcccaaggtcaaggaggg 1019

Qy    1561 gcaaccaactgccccagggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
           ||||||||||||||  ||||     ||  |   |||||||||||||||| |||||    |
Db    1020 gcaaccaactgcccagggggatatgggtcctcttggggccttcgggaccctggggncaag 1079

Qy    1621 gtgcttcc 1628
           | |||| |
Db    1080 gggctttc 1087


RESULT   5
US-09-971-429B-54/c
; Sequence 54, Application US/09971429B
; GENERAL INFORMATION:
;  APPLICANT: Lasek, Amy K. W.
;  APPLICANT:  Shyjan, Andrew W.
;  APPLICANT:  Turner, Christopher M.
;  TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
;  FILE REFERENCE: PA-0040 US
;  CURRENT APPLICATION NUMBER: US/09/971,429B
;  CURRENT FILING DATE:  2001-10-04
;  PRIOR APPLICATION NUMBER: 60/239,024
;  PRIOR FILING DATE: 2000-04-10
;  NUMBER OF SEQ ID NOS: 56
;  SOFTWARE: PERL Program
; SEQ ID NO 54
;   LENGTH: 1188
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 234340.19
US-09-971-429B-54


  Query Match             41.7%;  Score 790.6;  DB 5;  Length 1188;
  Best Local Similarity   98.9%;  Pred. No. 1.6e-140;
  Matches  796;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    1090 ctactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctc 1149
           | ||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     823 CCACCCAGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCATGCTC 764

Qy    1150 cgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagcta 1209
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     763 CGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCACAGCTA 704

Qy    1210 tggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagc 1269
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     703 TGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGAGGACTCAGC 644

Qy    1270 ccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagct 1329
           |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     643 CCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTACGACAAGCT 584

Qy    1330 gagccgctccatccgccagtcttacaagaagggcatcatccggaagccagacatctccca 1389
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     583 GAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGACATCTCCCA 524

Qy    1390 gcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaaacccg 1449
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    523 GCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCCTGAAACCCG 464

Qy   1450 ccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatgggggaaaac 1509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    463 CCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAAAC 404

Qy   1510 gggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggggcaaccaac 1569
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 GGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCAAC 344

Qy   1570 tgccccagggggatatgggtcctctggggccttcgggaccatggggcaggggtgcttcct 1629
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    343 TGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGGGTGCTTCCT 284

Qy   1630 cctcaggcccagctgctcccctggaggacagagggagacagggctgctccccaacacctg 1689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    283 CCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACCTG 224

Qy   1690 cctctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaaaggcc 1749
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223 CCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAAAGGCC 164

Qy   1750 acaggcagtccaggcctctctctgctccatccccctgcctcccattctgcaccacacctg 1809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    163 ACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCACCACACCTG 104

Qy   1810 gcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgccccccgggaat 1869
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    103 GCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCCCCCGGGAAT 44

Qy   1870 ggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||
Db     43 GGATAATAAAGATACTAGAGAACTG 19


RESULT   6
US-09-053-375B-626
; Sequence 626, Application US/09053375B
; GENERAL INFORMATION:
;  APPLICANT: Chenchik, Alex
;  APPLICANT:  Bibilashvilli, Robert
;  TITLE OF INVENTION: Nucleic Acid Arrays
;  FILE REFERENCE: CLON-006
;  CURRENT APPLICATION NUMBER: US/09/053,375B
;  CURRENT FILING DATE:  1998-08-31
;  NUMBER OF SEQ ID NOS: 1543
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 626
;   LENGTH: 2286
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-053-375B-626


  Query Match             4.2%;  Score 80.4;  DB 5;  Length 2286;
  Best Local Similarity   58.4%;  Pred. No. 2.6e-06;
  Matches  160;  Conservative    0;  Mismatches  111;  Indels    3;  Gaps    1;

Qy   1189 gctactcaagccccacagctatggccgcttcattaggtggctcaacaaggagaagggcat 1248
          ||| || |     ||  |  |||||| | ||||    |||  |     |||    ||
Db    512 gcttctgagagaacaaggtaatggccacatcatctcctggacctcacgggatggtggtga 571

Qy   1249 cttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcc 1308
           |||||  | | |||  |||   ||||||||||||| |||||  | |||||||||    |
Db    572 gttcaagttggtggatgcagaggaggtggcccggctatggggactgcgcaagaacaagac 631

Qy   1309 cgccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcat 1368
          |  |||||| | |||||||||| |||||  || | ||  | |  ||  | |||   |||||
Db    632 caacatgaattacgacaagcttagccgggccttgcggtactactatgataagaatatcat 691

Qy   1369 ccggaagccagacatctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcc 1428
          ||| |||   |  | |  ||||    | |||||| |||| |||  |  |      ||  |
Db    692 ccgcaag---gtgagcggccagaagtttgtctacaagtttgtgtcctacccagaggttgc 748

Qy   1429 tggcccagggcctgaaacccgccctcaggggcct 1462
          ||       |||||  | |||| |    ||||
Db    749 agggtgctccactgaagactgcccaccccagcct 782


RESULT   7
PCT-US01-43704-2290/c
; Sequence 2290, Application PC/TUS0143704
; GENERAL INFORMATION:
;  APPLICANT: Corixa Corporation
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Chenault, Ruth A.
;  APPLICANT:  Meagher, Madelein Joy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.561PC
;  CURRENT APPLICATION NUMBER: PCT/US01/43704
;  CURRENT FILING DATE:  2001-11-19
;  NUMBER OF SEQ ID NOS: 2606
;  SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2290
;   LENGTH: 499
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US01-43704-2290


  Query Match             4.0%;  Score 75;  DB 1;  Length 499;
  Best Local Similarity   59.0%;  Pred. No. 2.8e-05;
  Matches  148;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    345 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 286

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    285 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 226

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || ||||||      || ||      |  ||||| || ||| ||||||||||  |
Db    225 CCAACTATGGGGCCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 166

Qy   1339 catccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||  ||| | ||  |    |||| | | |  |       || ||| |||||
Db    165 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 109

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db    108 CTACAAGTTTG 98


RESULT   8
PCT-US02-07826-74
; Sequence 74, Application PC/TUS0207826
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc. et al.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
;  TITLE OF INVENTION:  Assessment, Prevention, and Therapy of Ovarian Cancer
;  FILE REFERENCE: MRI-030PC
;  CURRENT APPLICATION NUMBER: PCT/US02/07826
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: 60/276,025
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/325,149
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/276,026
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/324,967
;  PRIOR FILING DATE: 2001/09/26
;  PRIOR APPLICATION NUMBER: 60/311,732
```

```
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: 60/325,102
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/323,580
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
;   LENGTH: 1907
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-07826-74


  Query Match             4.0%;  Score 75;  DB 1;  Length 1907;
  Best Local Similarity   59.0%;  Pred. No. 2.8e-05;
  Matches  148;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    911 cacccacctgtgggagttcatccgggacatcctcatccacccggagctcaacgagggcct 970

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    971 catgaagtgggagaatcggcatgaaggcgtcttcaagttcctgcgctccgaggctgtggc 1030

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || ||||||      || ||      |  ||||| || ||| ||||||||||  |
Db   1031 ccaactatggggccaaaagaaaaagaacagcaacatgacctacgagaagctgagccgggc 1090

Qy   1339 catccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||  ||| | ||  |    |||| | | |  |       || ||| |||||
Db   1091 catgag---gtactactacaaacgggagatcctggaacgggtggatggccggcgactcgt 1147

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1148 ctacaagtttg 1158


RESULT   9
US-10-097-340-74
; Sequence 74, Application US/10097340
; GENERAL INFORMATION:
;  APPLICANT: John MONAHAN
;  APPLICANT:  Manjula GANNAVARAPU
;  APPLICANT:  Sebastian HOERSCH
;  APPLICANT:  Shubhangi KAMATKAR
;  APPLICANT:  Steve G. KOVATS
;  APPLICANT:  Rachel E. MEYERS
;  APPLICANT:  Michael MORRISEY
;  APPLICANT:  Peter OLANDT
;  APPLICANT:  Ami SEN
;  APPLICANT:  Peter VEIBY
;  APPLICANT:  Gordon B. MILLS
;  APPLICANT:  Robert C. BAST, Jr.
;  APPLICANT:  Karen LU
;  APPLICANT:  Rosemarie SCHMANDT
;  APPLICANT:  Xumei ZHAO
;  APPLICANT:  Karen GLATT
;  TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
;  TITLE OF INVENTION:  Assessment, Prevention, and Therapy of Ovarian Cancer
;  FILE REFERENCE: MRI-030
;  CURRENT APPLICATION NUMBER: US/10/097,340
;  CURRENT FILING DATE:  2002-03-14
;  PRIOR APPLICATION NUMBER: 60/276,025
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/325,149
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/276,026
;  PRIOR FILING DATE: 2001-03-14
;  PRIOR APPLICATION NUMBER: 60/324,967
;  PRIOR FILING DATE: 2001/09/26
;  PRIOR APPLICATION NUMBER: 60/311,732
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: 60/325,102
;  PRIOR FILING DATE: 2001-09-26
;  PRIOR APPLICATION NUMBER: 60/323,580
;  PRIOR FILING DATE: 2001-09-19
;  NUMBER OF SEQ ID NOS: 363
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74
;   LENGTH: 1907
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-097-340-74


  Query Match             4.0%;  Score 75;  DB 7;  Length 1907;
  Best Local Similarity   59.0%;  Pred. No. 2.8e-05;
  Matches  148;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    911 cacccacctgtgggagttcatccgggacatcctcatccacccggagctcaacgagggcct 970

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    971 catgaagtgggagaatcggcatgaaggcgtcttcaagttcctgcgctccgaggctgtggc 1030

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || ||||||      || ||      |  ||||| || ||| ||||||||||  |
Db   1031 ccaactatggggccaaaagaaaaagaacagcaacatgacctacgagaagctgagccgggc 1090

Qy   1339 catccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||  ||| | ||  |    |||| | | |  |       || ||| |||||
Db   1091 catgag---gtactactacaaacgggagatcctggaacgggtggatggccggcgactcgt 1147

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1148 ctacaagtttg 1158


RESULT  10
US-09-919-002-13047
; Sequence 13047, Application US/09919002
; GENERAL INFORMATION:
;  APPLICANT: Leshkowitz, Dena
;  APPLICANT:  Liu, Jin
;  TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
;  TITLE OF INVENTION:  LIBRARIES
;  FILE REFERENCE: 20411-752CON1
;  CURRENT APPLICATION NUMBER: US/09/919,002
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
;  PRIOR FILING DATE: FILING DATE: 1999-07-22
;  PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
;  PRIOR FILING DATE: FILING DATE: 1998-02-13
;  NUMBER OF SEQ ID NOS: 13203
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13047
;   LENGTH: 2457
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-919-002-13047


  Query Match             4.0%;  Score 75;  DB 5;  Length 2457;
  Best Local Similarity   59.0%;  Pred. No. 2.8e-05;
  Matches  148;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    940 cacccacctgtgggagttcatccgggacatcctcatccacccggagctcaacgagggcct 999

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
```

```
Db   1000 catgaagtgggagaatcggcatgaaggcgtcttcaagttcctgcgctccgaggctgtggc 1059

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || ||||||      || ||      |  ||||| || ||| ||||||||||  |
Db   1060 ccaactatggggccaaaagaaaaagaacagcaacatgacctacgagaagctgagccgggc 1119

Qy   1339 catccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||  ||| | ||  |    |||| | | |  |        || ||| |||||
Db   1120 catgag---gtactactacaaacgggagatcctggaacgggtggatggccggcgactcgt 1176

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1177 ctacaagtttg 1187


RESULT 11
US-10-099-926-1680
; Sequence 1680, Application US/10099926
; GENERAL INFORMATION:
;  APPLICANT: King, Gordon E.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.547C2
;  CURRENT APPLICATION NUMBER: US/10/099,926
;  CURRENT FILING DATE:  2002-03-17
;  NUMBER OF SEQ ID NOS: 1982
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1680
;   LENGTH: 2642
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-099-926-1680


  Query Match             4.0%;  Score 75;  DB 7;  Length 2642;
  Best Local Similarity   81.3%;  Pred. No. 2.8e-05;
  Matches   87;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1269 cccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagc 1328
          ||||||| ||| |||| |||||||||||  |||||||| || |||||||| | ||||||||
Db   1245 cccaggtcgccaggctctggggcatccagaagaaccggccagccatgaattacgacaagc 1304

Qy   1329 tgagccgctccatccgccagtcttacaagaagggcatcatccggaag 1375
          ||||||||||  ||||  | | |||  |||| ||||||||| | ||||
Db   1305 tgagccgctcgctccgatactattatgagaaaggcatcatgcagaag 1351


RESULT 12
US-09-654-788A-107
; Sequence 107, Application US/09654788A
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  OVARIAN CANCER
;  FILE REFERENCE: MRI-006A
;  CURRENT APPLICATION NUMBER: US/09/654,788A
;  CURRENT FILING DATE:  2000-09-01
;  PRIOR APPLICATION NUMBER: 60/152,547
;  PRIOR FILING DATE: 1999-09-03
;  PRIOR APPLICATION NUMBER: 60/190,347
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/191,321
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/208,382
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 60/220,467
;  PRIOR FILING DATE: 2000-07-20
;  NUMBER OF SEQ ID NOS: 157
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107
;   LENGTH: 3045
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-654-788A-107


  Query Match             4.0%;  Score 75;  DB 5;  Length 3045;
  Best Local Similarity   59.0%;  Pred. No. 2.8e-05;
  Matches  148;  Conservative    0;  Mismatches  100;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    982 cacccacctgtgggagttcatccgggacatcctcatccacccggagctcaacgagggcct 1041

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db   1042 catgaagtgggagaatcggcatgaaggcgtcttcaagttcctgcgctccgaggctgtggc 1101

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || ||||||      || ||      |  ||||| || ||| ||||||||||  |
Db   1102 ccaactatggggccaaaagaaaaagaacagcaacatgacctacgagaagctgagccgggc 1161

Qy   1339 catccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||  ||| | ||  |    |||| | | |  |        || ||| |||||
Db   1162 catgag---gtactactacaaacgggagatcctggaacgggtggatggccggcgactcgt 1218

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1219 ctacaagtttg 1229


RESULT 13
US-10-167-485-1
; Sequence 1, Application US/10167485
;   GENERAL INFORMATION:
;        APPLICANT: FUJINAGA, Kei
;                   YOSHIDA, Koichi
;                   HIGASHINO, Fumihiro
;        TITLE OF INVENTION: CANCER CONTROL
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
;             STREET: 624 Ninth Street N.W., Ste. 300
;             CITY: Washington
;             STATE: D.C.
;             COUNTRY: USA
;             ZIP: 20001
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/167,485
;             FILING DATE: 13-Jun-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/602,868
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: 08/875,944
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: PCT/JP96/00016
;             FILING DATE: 09-JAN-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BROWDY, Roger L.
;             REGISTRATION NUMBER: 25,618
;             REFERENCE/DOCKET NUMBER: FUJINAGA=1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202) 628-5197
```

```
;            TELEFAX: (202) 737-3528
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 2064 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..1386
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-167-485-1


  Query Match             3.9%;  Score 74.8;  DB 6;  Length 2064;
  Best Local Similarity   74.6%;  Pred. No. 3e-05;
  Matches   94;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
          ||||| |    ||  | |   |||| ||| |||| |||||||||| ||||||||  ||
Db   1039 TTCAAGCTCATTGAGCCTGAGGAGGTCGCCAGGCTCTGGGGCATCCAGAAGAACCGGCCA 1098

Qy   1310 gccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcatc 1369
          |||||||| | |||||||||||||||||  |||| | | |||  |||| ||||||||
Db   1099 GCCATGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAAAGGCATCATG 1158

Qy   1370 cggaag 1375
          | ||||
Db   1159 CAGAAG 1164


RESULT  14
US-10-099-926-1788
; Sequence 1788, Application US/10099926
; GENERAL INFORMATION:
;  APPLICANT: King, Gordon E.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.547C2
;  CURRENT APPLICATION NUMBER: US/10/099,926
;  CURRENT FILING DATE:  2002-03-17
;  NUMBER OF SEQ ID NOS: 1982
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1788
;   LENGTH: 2333
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-099-926-1788


  Query Match             3.9%;  Score 74.8;  DB 7;  Length 2333;
  Best Local Similarity   74.6%;  Pred. No. 3e-05;
  Matches   94;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
          ||||| |    ||  | |   |||| ||| |||| |||||||||| ||||||||  ||
Db   1308 ttcaagctcattgagcctgaggaggtcgccaggctctggggcatccagaagaaccggcca 1367

Qy   1310 gccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcatc 1369
          |||||||| | |||||||||||||||||  |||| | | |||  |||| ||||||||
Db   1368 gccatgaattacgacaagctgagccgctcgctccgatactattatgagaaaggcatcatg 1427

Qy   1370 cggaag 1375
          | ||||
Db   1428 cagaag 1433


RESULT  15
PCT-US02-12378-282
; Sequence 282, Application PC/TUS0212378
; GENERAL INFORMATION:
;  APPLICANT: Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Jiang, Yuqiu
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Durham, Margarita
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF BREAST CANCER
;  FILE REFERENCE: 210121.47003PC
;  CURRENT APPLICATION NUMBER: PCT/US02/12378
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 627
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 282
;   LENGTH: 502
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US02-12378-282


  Query Match             3.9%;  Score 74;  DB 1;  Length 502;
  Best Local Similarity   59.5%;  Pred. No. 4.3e-05;
  Matches  125;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    269 cacccacctgtgggagttcatccgggacatcctcatccacccggagctcaacgagggcct 328

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    329 catgaagtgggagaatcggcatgaaggcgtcttcaagttcctgcgctccgaggctgtggc 388

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || ||||||      || ||      |  ||||| || ||| ||||||||||  |
Db    389 ccaactatggggccaaaagaaaaagaacagcaacatgacctacgagaagctgagccgggc 448

Qy   1339 catccgccagtcttacaagaagggcatcat 1368
          |||  |  | |  |||||   ||  ||| |
Db    449 catgaggtactactacaaacgggagatcct 478


Search completed: August  6, 2002, 20:55:39
Job time: 9463 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:16:51 ; Search time 87.61 Seconds
(without alignments)
5310.237 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1861.4	98.3	1905	4	US-09-055-113-2	Sequence 2, Appli
	2	263	13.9	278	4	US-09-439-313-398	Sequence 398, App
	3	83.2	4.4	2266	2	US-09-213-767-1	Sequence 1, Appli
	4	75.2	4.0	2667	2	US-08-469-412A-1	Sequence 1, Appli
	5	75.2	4.0	2667	4	US-09-021-715-1	Sequence 1, Appli
	6	74.8	3.9	2064	3	US-08-875-944B-1	Sequence 1, Appli
	7	74.8	3.9	2064	4	US-09-116-049-3	Sequence 3, Appli
	8	73.4	3.9	1920	1	US-08-746-789A-1	Sequence 1, Appli
	9	71	3.7	2938	2	US-08-343-443B-3	Sequence 3, Appli
	10	70.2	3.7	2410	2	US-08-780-835B-1	Sequence 1, Appli
	11	70.2	3.7	2410	4	US-09-303-268-1	Sequence 1, Appli
	12	70.2	3.7	2410	4	US-09-116-049-1	Sequence 1, Appli
	13	64.8	3.4	1752	4	US-09-360-779-1	Sequence 1, Appli
	14	63	3.3	2268	3	US-09-344-579-1	Sequence 1, Appli
	15	58.4	3.1	2544	2	US-08-469-412A-6	Sequence 6, Appli
	16	58.4	3.1	2544	4	US-09-021-715-6	Sequence 6, Appli
c	17	57.4	3.0	852	4	US-09-020-956-44	Sequence 44, Appl
c	18	57.4	3.0	852	4	US-09-030-607-44	Sequence 44, Appl
c	19	57.4	3.0	852	4	US-09-439-313-44	Sequence 44, Appl
	20	57.4	3.0	5427	3	US-09-009-913-2	Sequence 2, Appli
	21	57.4	3.0	5510	3	US-09-009-913-3	Sequence 3, Appli
	22	57.4	3.0	5667	3	US-09-009-913-4	Sequence 4, Appli
	23	57.2	3.0	2975	1	US-08-368-281-1	Sequence 1, Appli
	24	57.2	3.0	3240	1	US-08-368-281-3	Sequence 3, Appli
	25	56.4	3.0	1604	1	US-08-306-691B-43	Sequence 43, Appl
	26	56.4	3.0	1604	5	PCT-US93-06251-9	Sequence 9, Appli
	27	54.8	2.9	848	3	US-09-009-913-338	Sequence 338, App
	28	50.2	2.7	1447	4	US-08-878-177-1	Sequence 1, Appli
	29	50.2	2.7	1528	4	US-08-878-177-3	Sequence 3, Appli
	30	46.2	2.4	1035	1	US-07-601-094-30	Sequence 30, Appl
	31	46.2	2.4	1035	1	US-08-012-735-30	Sequence 30, Appl
	32	46.2	2.4	1281	4	US-08-998-416-3	Sequence 3, Appli
	33	46.2	2.4	1910	4	US-09-593-711A-3	Sequence 3, Appli
	34	46.2	2.4	1914	1	US-07-601-094-1	Sequence 1, Appli
	35	46.2	2.4	1914	1	US-08-012-735-1	Sequence 1, Appli
	36	45.8	2.4	1364	1	US-08-306-691B-50	Sequence 50, Appl
	37	45.8	2.4	1364	5	PCT-US93-06251-65	Sequence 65, Appl
	38	45.2	2.4	152331	3	US-09-128-155-16	Sequence 16, Appl
c	39	44.6	2.4	985	4	US-09-056-556-182	Sequence 182, App
	40	44	2.3	23673	4	US-09-773-816-1	Sequence 1, Appli
	41	43.6	2.3	15872	4	US-09-105-537-1	Sequence 1, Appli
	42	43.6	2.3	43280	2	US-08-804-227C-1	Sequence 1, Appli
	43	43	2.3	19182	2	US-08-850-880-11	Sequence 11, Appl
	44	43	2.3	19182	2	US-08-944-916-11	Sequence 11, Appl
	45	42.6	2.2	1227	4	US-09-385-028-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-055-113-2
; Sequence 2, Application US/09055113
; Patent No. 6265565
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: PROSTATE-ASSOCIATED ETS PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,113
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0501 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSTUT12
; CLONE: 1813005
US-09-055-113-2

Query Match 98.3%; Score 1861.4; DB 4; Length 1905;

```
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1871;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 74

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 134

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 194

Qy     181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
           |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 254

Qy     241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 314

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     315 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 374

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 434

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 494

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 554

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     555 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 614

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 674

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db     675 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 734

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 794

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 854

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     855 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 914

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     915 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 974

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     975 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1034

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    1035 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1094

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1095 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1154

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1155 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1214

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1215 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1274

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1275 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1334

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1335 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1394

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1395 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1454

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1455 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1514

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1515 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1574

Qy    1561 gcaaccaactgccccaggggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    1575 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGG 1634

Qy    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1635 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1694

Qy    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1695 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1754

Qy    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1755 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1814

Qy    1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    1815 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1874

Qy    1861 cccgggaatggataataaagatactag 1887
           |||||||||||||||||||||||||||
Db    1875 CCCGGGAATGGATAATAAAGATACTAG 1901


RESULT   2
US-09-439-313-398
; Sequence 398, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan Louise
;  APPLICANT:  Jiang Yuqui
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Kalos, Michael
;  APPLICANT:  Fanger, Gary
;  APPLICANT:  Retter, Mark
```

```
;  APPLICANT:  Solk, John
;  APPLICANT:  Day, Craig
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C9
;  CURRENT APPLICATION NUMBER: US/09/439,313
;  CURRENT FILING DATE:  1999-11-12
;  NUMBER OF SEQ ID NOS: 575
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 398
;   LENGTH: 278
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(278)
;   OTHER INFORMATION: n = A,T,C or G
US-09-439-313-398


  Query Match             13.9%;  Score 263;  DB 4;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 3.8e-48;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 74

Qy    1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 134

Qy    1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 194

Qy    1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 254

Qy    1221 ttaggtggctcaacaaggagaagg 1244
           ||| ||||||||||||||||||||
Db     255 ttangtggctcaacaaggagaagg 278


RESULT   3
US-09-213-767-1
; Sequence 1, Application US/09213767
; Patent No. 5948680
; GENERAL INFORMATION:
;  APPLICANT: Brenda F. Baker
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF ELK-1 EXPRESSION
;  FILE REFERENCE: RTS-0024
;  CURRENT APPLICATION NUMBER: US/09/213,767
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
;   LENGTH: 2266
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (316)..(1602)
US-09-213-767-1


  Query Match             4.4%;  Score 83.2;  DB 2;  Length 2266;
  Best Local Similarity   59.8%;  Pred. No. 2.5e-09;
  Matches  177;  Conservative    0;  Mismatches  113;  Indels    6;  Gaps    2;

Qy    1166 ctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcattagg 1225
           |||||||||||| ||   | || ||||     ||   ||  || |||||| | ||||
Db     334 ctgtggcagtttctgctgcagctgctgagagag---caaggcaatggccacatcatctcc 390

Qy    1226 tggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccggctg 1285
           |||         |||    ||    |||||  | | |||  |||   ||||||||||||||
Db     391 tggacttcacgggatggtggtgaattcaagctggtggatgcagaggaggtggcccggctg 450

Qy    1286 tggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgc 1345
           |||||  | |||||||||    ||  |||||| | |||||||| |||||  || | ||
Db     451 tggggactacgcaagaacaagaccaacatgaattacgacaagctcagccgggccttgcgg 510

Qy    1346 cagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgtctaccag 1405
           | |  ||  | |||  |||||||||| |||   |  | |  ||||    |||||||| ||
Db     511 tactactatgacaagaacatcatccgcaag---gtgagcggccagaagttcgtctacaag 567

Qy    1406 ttcgtgcaccccatctgagtgcctggcccagggcctgaaacccgccctcaggggcc 1461
           || |||  |  |      ||  | ||        ||||   | |||| |    |||
Db     568 tttgtgtcctaccctgaggtcgcagggtgctccactgaggactgcccgccccagcc 623


RESULT   4
US-08-469-412A-1
; Sequence 1, Application US/08469412A
; Patent No. 5856125
;  GENERAL INFORMATION:
;    APPLICANT:  Mavrothalassitis, George J.
;    APPLICANT:  Blair, Donald G.
;    APPLICANT:  Fisher, Robert J.
;    APPLICANT:  Beal Jr., Gregory J.
;    APPLICANT:  Athanasiou, Meropi A.
;    APPLICANT:  Sgouras, Dionyssios N.
;    TITLE OF INVENTION:  The ERF Genetic Locus and Its Products
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,412A
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  015280-229000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2667 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  123..1769
;      OTHER INFORMATION:  /note= "human ERF (ETS2 Repressor
;      OTHER INFORMATION:  Factor) cDNA"
US-08-469-412A-1


  Query Match             4.0%;  Score 75.2;  DB 2;  Length 2667;
  Best Local Similarity   59.1%;  Pred. No. 1.3e-07;
  Matches  150;  Conservative    0;  Mismatches   98;  Indels    6;  Gaps    1;
```

```
Qy   1124 gacagcgaggtggactcatcatgctccgggcagcccatccacctgtggcagttcctcaag 1183
          | |  | ||   || || ||       |      ||  ||||| |||||||| ||  ||  |
Db    165 GCCTACAAGCCAGAGTCGTCCCCTGGCTCAAGGCAGATCCAGCTGTGGCACTTTATCCTG 224

Qy   1184 gagttgctactcaagccccacagctatggccgcttcattaggtggctcaacaaggagaag 1243
          ||| |||| |  |||    |   | | |||  | |     ||  |    || | | | | |
Db    225 GAGCTGCTGCGGAAGGAGGAGTACCAGGGCGTCATTGCCTGGCAGGGGGACTACGGGGAA 284

Qy   1244 ggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaac 1303
            | || |||||      ||| | |   ||||||||||||||||||| | ||||||  |
Db    285 TTCGTCATCAAA------GACCCTGATGAGGTGGCCCGGCTGTGGGGCGTTCGCAAGTGC 338

Qy   1304 cgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggc 1363
             |||   ||||| | ||||||||||||||  || | ||| | |  || || ||| ||
Db    339 AAGCCCCAGATGAATTACGACAAGCTGAGCCGGGCCCTGCGCTATTACTATAACAAGCGC 398

Qy   1364 atcatccggaagcc 1377
          ||  | |  ||| |
Db    399 ATTCTGCACAAGAC 412
```

```
RESULT   5
US-09-021-715-1
; Sequence 1, Application US/09021715
; Patent No. 6194547
;   GENERAL INFORMATION:
;        APPLICANT: Mavrothalassitis, George J.
;                   Blair, Donald G.
;                   Fisher, Robert J.
;                   Beal Jr., Gregory J.
;                   Athanasiou, Meropi A.
;                   Sgouras, Dionyssios N.
;        TITLE OF INVENTION: The ERF Genetic Locus and Its Products
;        NUMBER OF SEQUENCES: 16
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/021,715
;             FILING DATE: 10-Feb-1998
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Garrett-Wackowski, Eugenia
;             REGISTRATION NUMBER: 37,330
;             REFERENCE/DOCKET NUMBER: 015280-229000
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2667 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  123..1769
;             OTHER INFORMATION:  /note= "human ERF (ETS2 Repressor
;             Factor) cDNA"
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-021-715-1
```

```
  Query Match             4.0%;  Score 75.2;  DB 4;  Length 2667;
  Best Local Similarity   59.1%;  Pred. No. 1.3e-07;
  Matches  150;  Conservative    0;  Mismatches   98;  Indels    6;  Gaps    1;

Qy   1124 gacagcgaggtggactcatcatgctccgggcagcccatccacctgtggcagttcctcaag 1183
          | |  | ||   || || ||       |      ||  ||||| |||||||| ||  ||  |
Db    165 GCCTACAAGCCAGAGTCGTCCCCTGGCTCAAGGCAGATCCAGCTGTGGCACTTTATCCTG 224

Qy   1184 gagttgctactcaagccccacagctatggccgcttcattaggtggctcaacaaggagaag 1243
          ||| |||| |  |||    |   | | |||  | |     ||  |    || | | | | |
Db    225 GAGCTGCTGCGGAAGGAGGAGTACCAGGGCGTCATTGCCTGGCAGGGGGACTACGGGGAA 284

Qy   1244 ggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaac 1303
            | || |||||      ||| | |   ||||||||||||||||||| | ||||||  |
Db    285 TTCGTCATCAAA------GACCCTGATGAGGTGGCCCGGCTGTGGGGCGTTCGCAAGTGC 338

Qy   1304 cgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggc 1363
             |||   ||||| | ||||||||||||||  || | ||| | |  || || ||| ||
Db    339 AAGCCCCAGATGAATTACGACAAGCTGAGCCGGGCCCTGCGCTATTACTATAACAAGCGC 398

Qy   1364 atcatccggaagcc 1377
          ||  | |  ||| |
Db    399 ATTCTGCACAAGAC 412
```

```
RESULT   6
US-08-875-944B-1
; Sequence 1, Application US/08875944B
; Patent No. 6096542
;  GENERAL INFORMATION:
;    APPLICANT:  FUJINAGA, Kei
;    APPLICANT:  YOSHIDA, Koichi
;    APPLICANT:  HIGASHINO, Fumihiro
;    TITLE OF INVENTION:  CANCER CONTROL
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK, P.L.L.C.
;      STREET:  624 Ninth Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/875,944B
;      FILING DATE:  07-AUG-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 07-020173
;      FILING DATE:  08-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP96/00016
;      FILING DATE:  09-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROWDY, Roger L.
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  FUJINAGA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2064 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
```

```
;      NAME/KEY:  CDS
;      LOCATION:  1..1386
US-08-875-944B-1


  Query Match             3.9%;  Score 74.8;  DB 3;  Length 2064;
  Best Local Similarity   74.6%;  Pred. No. 1.5e-07;
  Matches   94;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
          |||||  |    ||  | |   |||| ||| |||| ||||||||||  |||||||| ||
Db   1039 TTCAAGCTCATTGAGCCTGAGGAGGTCGCCAGGCTCTGGGGCATCCAGAAGAACCGGCCA 1098

Qy   1310 gccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcatc 1369
          |||||||| | ||||||||||||||||||  ||||  | | |||  |||| ||||||||
Db   1099 GCCATGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAAAGGCATCATG 1158

Qy   1370 cggaag 1375
          | ||||
Db   1159 CAGAAG 1164


RESULT   7
US-09-116-049-3
; Sequence 3, Application US/09116049A
; Patent No. 6248351
; GENERAL INFORMATION:
;  APPLICANT: Hung, Men-Chie
;  TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
;  FILE REFERENCE: UTSC:582
;  CURRENT APPLICATION NUMBER: US/09/116,049A
;  CURRENT FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2064
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-116-049-3


  Query Match             3.9%;  Score 74.8;  DB 4;  Length 2064;
  Best Local Similarity   74.6%;  Pred. No. 1.5e-07;
  Matches   94;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
          |||||  |    ||  | |   |||| ||| |||| ||||||||||  |||||||| ||
Db   1039 ttcaagctcattgagcctgaggaggtcgccaggctctggggcatccagaagaaccggcca 1098

Qy   1310 gccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcatc 1369
          |||||||| | ||||||||||||||||||  ||||  | | |||  |||| ||||||||
Db   1099 gccatgaattacgacaagctgagccgctcgctccgatactattatgagaaaggcatcatg 1158

Qy   1370 cggaag 1375
          | ||||
Db   1159 cagaag 1164


RESULT   8
US-08-746-789A-1
; Sequence 1, Application US/08746789A
; Patent No. 5789200
;  GENERAL INFORMATION:
;    APPLICANT:  Ismail Kola, Martin J. Tymms, Christine DeBouck
;    TITLE OF INVENTION:  A No. 5789200el Human ETS Family Member, ELF3
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road, P.O. Box 1539
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  MICROSOFT WORD
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/746,789A
;      FILING DATE:  No. 5789200ember 15, 1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  William T. Han
;      REGISTRATION NUMBER:  34,344
;      REFERENCE/DOCKET NUMBER:  ATG 50024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610 270 5219
;      TELEFAX:  610 270 4026
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1920
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
US-08-746-789A-1


  Query Match             3.9%;  Score 73.4;  DB 1;  Length 1920;
  Best Local Similarity   58.6%;  Pred. No. 3e-07;
  Matches  147;  Conservative    0;  Mismatches  101;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    930 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 989

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    990 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 1049

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctc 1338
          ||  || |||||       || ||      |  ||||| || ||| ||||||||||  |
Db   1050 CCAACTATGGGGGCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 1109

Qy   1339 catccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||  ||| | ||  |    |||| | | |  |       || ||| |||||
Db   1110 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1166

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1167 CTACAAGTTTG 1177


RESULT   9
US-08-343-443B-3
; Sequence 3, Application US/08343443B
; Patent No. 5968734
;  GENERAL INFORMATION:
;    APPLICANT:  Aurias, Alain
;    APPLICANT:  Delattre, Olivier
;    APPLICANT:  Desmaze, Chantal
;    APPLICANT:  Melot, Thomas
;    APPLICANT:  Peter, Martine
;    APPLICANT:  Ploougastel, Beatrice
;    APPLICANT:  Thomas, Gilles
;    APPLICANT:  Zucman, Jessica
;    TITLE OF INVENTION:  NUCLEIC ACID CORRESPONDING TO A GENE OF
;    TITLE OF INVENTION:  CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
;    TITLE OF INVENTION:  TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
;    TITLE OF INVENTION:  TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
;    TITLE OF INVENTION:  TRANSLOCATIONS
;    NUMBER OF SEQUENCES:  129
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates
;      STREET:  230 South Fifteenth Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  AEDIT 1.0 DOS text editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/343,443B
;      FILING DATE:  18-NOV-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00494
;      FILING DATE:  19-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92/06123
;      FILING DATE:  20-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiser, Gerard J.
;      REGISTRATION NUMBER:  19,763
;      REFERENCE/DOCKET NUMBER:  989.6121P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-875-8383
;      TELEFAX:  215-875-8394
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2938 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  143..1498
US-08-343-443B-3


  Query Match             3.7%;  Score 71;  DB 2;  Length 2938;
  Best Local Similarity   62.0%;  Pred. No. 1.1e-06;
  Matches  132;  Conservative    0;  Mismatches   75;  Indels    6;  Gaps    1;

Qy   1156 gcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccg 1215
          ||  ||||| |||||||| ||||||  |||| ||||   | |   |  || |    || |
Db    979 GCAGATCCAGCTGTGGCAATTCCTCCTGGAGCTGCTCTCCGACAGCGCCAACGCCAGCTG 1038

Qy   1216 cttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggt 1275
            |||   ||  |   | ||| | | ||      ||||||||   |||| | |   ||||
Db   1039 TATCACCTGGGAGGGGACCAACGGGGAG------TTCAAAATGACGGACCCCGATGAGGT 1092

Qy   1276 ggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccg 1335
          |||| |||  ||||||   || || | |   |||  |||||| | ||||||||||||||
Db   1093 GGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAACATGAATTACGACAAGCTGAGCCG 1152

Qy   1336 ctccatccgccagtcttacaagaagggcatcat 1368
            || ||||  | |  ||  | ||   ||| ||
Db   1153 GGCCCTCCGTTATTACTATGATAAAAACATTAT 1185


RESULT  10
US-08-780-835B-1
; Sequence 1, Application US/08780835B
; Patent No. 5922688
;  GENERAL INFORMATION:
;    APPLICANT:  Hung, Mien-Chie
;    APPLICANT:  Xing, Xiangming
;    TITLE OF INVENTION:  PEA3 is a Tumor Suppressor
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ARNOLD, WHITE AND DURKEE
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/780,835B
;      FILING DATE:  10-JAN-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mark B.
;      REGISTRATION NUMBER:  37,259
;      REFERENCE/DOCKET NUMBER:  UTSC500
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2410 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-780-835B-1


  Query Match             3.7%;  Score 70.2;  DB 2;  Length 2410;
  Best Local Similarity   59.2%;  Pred. No. 1.6e-06;
  Matches  158;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    2;

Qy   1109 gaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctg 1168
          || | ||   |  |   | | |||| |    | |       | ||    || | || |||
Db   1308 GAAGGGATTGGAGCTTTCCGGGAGGGGCCACCCTACCAGCGCCGGGGTGCCTTACAACTG 1367

Qy   1169 tggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtgg 1228
          |||||||| ||   || |   || ||  |   ||  |   |||  |  ||||||   |||
Db   1368 TGGCAGTTTCT---GGTGGCCCTGCTGGATGACCCAACAAATGCTCATTTCATTGCTTGG 1424

Qy   1229 ctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtgg 1288
            |  ||  | ||     || ||| |   ||  | |   |||| ||| |||| |||
Db   1425 ACAGGCCGGGGAATGG---AGTTTAAACTAATTGAACCTGAAGAGGTTGCCAGGCTCTGG 1481

Qy   1289 ggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccag 1348
          || ||||  |||||||| || |||||||| |  |||||||||||||||  | ||  |
Db   1482 GGTATCCAGAAGAACCGGCCAGCCATGAATTATGACAAGCTGAGCCGCTCGCTGCGATAC 1541

Qy   1349 tcttacaagaagggcatcatccggaag 1375
          | |||  |||| |||||||| | ||||
Db   1542 TATTATGAGAAAGGCATCATGCAGAAG 1568


RESULT  11
US-09-303-268-1
; Sequence 1, Application US/09303268
; Patent No. 6172212
;  GENERAL INFORMATION:
;       APPLICANT: Hung, Mien-Chie
;                  Xing, Xiangming
;       TITLE OF INVENTION: PEA3 is a Tumor Suppressor
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: ARNOLD, WHITE AND DURKEE
;            STREET: P.O. Box 4433
;            CITY: Houston
;            STATE: Texas
;            COUNTRY: USA
;            ZIP: 77210-4433
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
```

```
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;          CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/303,268
;             FILING DATE: 30-Apr-1999
;             CLASSIFICATION: <Unknown>
;          PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/780,835
;             FILING DATE: 10-JAN-1997
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Wilson, Mark B.
;             REGISTRATION NUMBER: 37,259
;             REFERENCE/DOCKET NUMBER: UTSC500
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (512) 418-3000
;             TELEFAX: (512) 474-7577
;      INFORMATION FOR SEQ ID NO: 1:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 2410 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-303-268-1


  Query Match             3.7%;  Score 70.2;  DB 4;  Length 2410;
  Best Local Similarity   59.2%;  Pred. No. 1.6e-06;
  Matches  158;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    2;

Qy   1109 gaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctg 1168
          || | ||   |  |   | | |||| |    | |       | ||    || | || |||
Db   1308 GAAGGGATTGGAGCTTTCCGGGAGGGGCCACCCTACCAGCGCCGGGGTGCCTTACAACTG 1367

Qy   1169 tggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtgg 1228
          |||||||| ||   || |   || ||  |   ||  |   |||  | ||||||   |||
Db   1368 TGGCAGTTTCT---GGTGGCCCTGCTGGATGACCCAACAAATGCTCATTTCATTGCTTGG 1424

Qy   1229 ctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtgg 1288
               |  ||  | ||     || ||| |    ||  | |   |||| ||| |||| |||
Db   1425 ACAGGCCGGGGAATGG---AGTTTAAACTAATTGAACCTGAAGAGGTTGCCAGGCTCTGG 1481

Qy   1289 ggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccag 1348
          || ||||  |||||||| || |||||||| |  ||||||||||||||||  | ||  |
Db   1482 GGTATCCAGAAGAACCGGCCAGCCATGAATTATGACAAGCTGAGCCGCTCGCTGCGATAC 1541

Qy   1349 tcttacaagaagggcatcatccggaag 1375
          | |||  |||| ||||||||  | ||||
Db   1542 TATTATGAGAAAGGCATCATGCAGAAG 1568


RESULT  12
US-09-116-049-1
; Sequence 1, Application US/09116049A
; Patent No. 6248351
; GENERAL INFORMATION:
;  APPLICANT: Hung, Men-Chie
;  TITLE OF INVENTION: HUMAN PEA3 IS A TUMOR SUPPRESSOR FOR CANCER CELLS
;  FILE REFERENCE: UTSC:582
;  CURRENT APPLICATION NUMBER: US/09/116,049A
;  CURRENT FILING DATE: 1998-07-15
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2410
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-116-049-1


  Query Match             3.7%;  Score 70.2;  DB 4;  Length 2410;
  Best Local Similarity   59.2%;  Pred. No. 1.6e-06;
  Matches  158;  Conservative    0;  Mismatches  103;  Indels    6;  Gaps    2;

Qy   1109 gaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctg 1168
          || | ||   |  |   | | |||| |    | |       | ||    || | || |||
Db   1308 gaagggattggagctttccgggaggggccaccctaccagcgccggggtgccttacaactg 1367

Qy   1169 tggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtgg 1228
          |||||||| ||   || |   || ||  |   ||  |   |||  | ||||||   |||
Db   1368 tggcagtttct---ggtggccctgctggatgacccaacaaatgctcatttcattgcttgg 1424

Qy   1229 ctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtgg 1288
               |  ||  | ||     || ||| |    ||  | |   |||| ||| |||| |||
Db   1425 acaggccggggaatgg---agtttaaactaattgaacctgaagaggttgccaggctctgg 1481

Qy   1289 ggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccag 1348
          || ||||  |||||||| || |||||||| |  ||||||||||||||||  | ||  |
Db   1482 ggtatccagaagaaccggccagccatgaattatgacaagctgagccgctcgctgcgatac 1541

Qy   1349 tcttacaagaagggcatcatccggaag 1375
          | |||  |||| ||||||||  | ||||
Db   1542 tattatgagaaaggcatcatgcagaag 1568


RESULT  13
US-09-360-779-1
; Sequence 1, Application US/09360779
; Patent No. 6268216
; GENERAL INFORMATION:
;  APPLICANT: Deneris, Evan S.
;  APPLICANT: Fyodoro, Dmitry V.
;  APPLICANT: Hendricks, Timothy J.
;  TITLE OF INVENTION: Reagents and Methods for the Screening of Compounds
;  TITLE OF INVENTION: Useful in the Treatment of Neurological Diseases
;  FILE REFERENCE: CASE-03828
;  CURRENT APPLICATION NUMBER: US/09/360,779
;  CURRENT FILING DATE: 1999-07-26
;  EARLIER APPLICATION NUMBER: 60/094,264
;  EARLIER FILING DATE: 1998-07-27
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1752
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (112)..(1131)
US-09-360-779-1


  Query Match             3.4%;  Score 64.8;  DB 4;  Length 1752;
  Best Local Similarity   53.1%;  Pred. No. 2.1e-05;
  Matches  186;  Conservative    0;  Mismatches  157;  Indels    7;  Gaps    2;

Qy   1156 gcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccg 1215
          ||  |||||  |||||||||| ||   |||| ||||     | | |   | |   ||| |
Db    555 gcagatccagttgtggcagtttctactggagctgctggcagaccgcgcgaacgccggctg 614

Qy   1216 cttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggt 1275
          | ||    ||  | |       |   | |||   |||||  |    ||| | | | ||||
Db    615 catcgcgtgggagggc------ggccacggcgagttcaagctcaccgaccccgacgaggt 668

Qy   1276 ggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccg 1335
          ||| || |  ||||||   ||||||| |   |||   ||||||| |||||||| || ||
Db    669 ggcgcgacgctggggcgagcgcaagagcaagcccaatatgaactacgacaagctaagtcg 728

Qy   1336 ctccatccgccagtcttacaagaagggcatcatccggaagccagacatctcccagcgcct 1395
            |  ||| | |  ||| | ||   ||||||  | |||    ||  |   |    |
Db    729 agcactgcgctactactacgacaaaaacatcatgagcaaggtgcacggcaagcgctacgc 788

Qy   1396 cgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaaacccgccctca 1455
          |  |  |     | | ||  | |  |  | || |||    ||     | | | | ||
Db    789 ctaccgctttgacttccagggcctggcacaggcttgccagccaccacccgcgcacgccca 848
```

```
Qy   1456 ggggcctctctcctgcctgccctgcctcagccaggccctgagatgggggа 1505
          | |  | || | |||  | | ||  |||||    ||   ||||| | |
Db    849 -cgccgccgctgccgccgccgcagcggcagccgccgcccaggatggcgca 897


RESULT 14
US-09-344-579-1
; Sequence 1, Application US/09344579
; Patent No. 6054316
; GENERAL INFORMATION:
;  APPLICANT: Brenda F. Baker
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF ETS-2 EXPRESSION
;  FILE REFERENCE: RTS-0063
;  CURRENT APPLICATION NUMBER: US/09/344,579
;  CURRENT FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
;   LENGTH: 2268
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (291)..(1700)
US-09-344-579-1


  Query Match             3.3%;  Score 63;  DB 3;  Length 2268;
  Best Local Similarity   58.8%;  Pred. No. 5.4e-05;
  Matches  130;  Conservative    0;  Mismatches   85;  Indels    6;  Gaps    1;

Qy   1157 cccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgc 1216
          || || || |||||||||| |||  |||| |||||       |  |||  |   |||
Db   1374 cctattcagctgtggcagtttctcctggagctgcta------tcagacaaatcctgccag 1427

Qy   1217 ttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtg 1276
          |   | |   |    |    ||     |    || ||  | |  ||| | |   |||||
Db   1428 tcattcatcagctggactggagacggatgggagtttaagctcgccgaccccgatgaggtg 1487

Qy   1277 gcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaagctgagccgc 1336
          |||||| | |||||| |   | || ||    |||   ||||||| ||| ||||||||||
Db   1488 gcccgccggtggggaaagaggaaaaataagcccaagatgaactacgagaagctgagccgg 1547

Qy   1337 tccatccgccagtcttacaagaagggcatcatccggaagcc 1377
            | | ||| | | |||| | |||  ||||||||  ||| |
Db   1548 ggcttacgctactattacgacaagaacatcatccacaagac 1588


RESULT 15
US-08-469-412A-6
; Sequence 6, Application US/08469412A
; Patent No. 5856125
;  GENERAL INFORMATION:
;    APPLICANT:  Mavrothalassitis, George J.
;    APPLICANT:  Blair, Donald G.
;    APPLICANT:  Fisher, Robert J.
;    APPLICANT:  Beal Jr., Gregory J.
;    APPLICANT:  Athanasiou, Meropi A.
;    APPLICANT:  Sgouras, Dionyssios N.
;    TITLE OF INVENTION:  The ERF Genetic Locus and Its Products
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,412A
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  015280-229000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2544 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..2544
;      OTHER INFORMATION:  /note= "genomic DNA sequence for murine
;      OTHER INFORMATION:  ERF gene (promoter region and first exon
;      OTHER INFORMATION:  not included)"
US-08-469-412A-6


  Query Match             3.1%;  Score 58.4;  DB 2;  Length 2544;
  Best Local Similarity   58.6%;  Pred. No. 0.00054;
  Matches  123;  Conservative    0;  Mismatches   81;  Indels    6;  Gaps    1;

Qy   1135 ggactcatcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctact 1194
          ||| |||||      |     ||  ||||| |||||||| ||  ||  |||| |||| |
Db    416 GGAGTCATCCCCTGGCTCCAGGCAGATCCAGCTGTGGCACTTTATCCTGGAGCTGCTTCG 475

Qy   1195 caagccccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaa 1254
          ||     |   | | |||  | ||  | ||  |    || | | | ||    || |||
Db    476 GAAAGAGGAGTACCAGGGCGTCATCGCTTGGCAGGGGGACTACGGGGAGTTTGTCATCA- 534

Qy   1255 aattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccat 1314
               ||||| | |   | ||||| || || |||||  ||||||||  |   |||   ||
Db    535 -----AGGACCCTGATGAAGTGGCTCGCCTCTGGGGGGTCCGCAAGTGCAAACCCCAGAT 589

Qy   1315 gaactgcgacaagctgagccgctccatccg 1344
          |||||  |||||||||||||||  |  | ||
Db    590 GAACTATGACAAGCTGAGCCGGGCTTTGCG 619


Search completed: August  6, 2002, 19:52:32
Job time: 5741 sec
```

GenCore version 4.5 ,

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:16:11 ; Search time 3492.95 Seconds
(without alignments)
11347.102 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1870	98.7	1894	6	AX048148	AX048148 Sequence
2	1870	98.7	1894	9	AF071538	AF071538 Homo sapi
3	1868.4	98.6	1918	6	AX175516	AX175516 Sequence
4	1865.2	98.5	3317	6	AX048159	AX048159 Sequence
5	1864.4	98.4	1942	9	BC021299	BC021299 Homo sapi
6	1811.8	95.7	1898	9	AB031549	AB031549 Homo sapi
7	757	40.0	1856	10	BC012648	BC012648 Mus muscu
8	738.4	39.0	1704	10	AB019436	AB019436 Mus muscu
c 9	643	33.9	150660	9	AL157372	AL157372 Human DNA
10	329.8	17.4	1751	6	AX048155	AX048155 Sequence
11	263	13.9	278	6	AX106617	AX106617 Sequence
12	263	13.9	278	6	AX140908	AX140908 Sequence
13	263	13.9	278	6	AX200768	AX200768 Sequence
14	263	13.9	278	6	AX267424	AX267424 Sequence
15	156.6	8.3	105772	2	AC095858	AC095858 Rattus no
c 16	151.2	8.0	105772	2	AC095858	AC095858 Rattus no
17	147.6	7.8	362	3	DRODETS4A	M88474 Drosophila
18	147.6	7.8	38178	2	AC015374	AC015374 Drosophil
19	147.6	7.8	180887	3	AC008253	AC008253 Drosophil
20	147.6	7.8	188549	3	AC007818	AC007818 Drosophil
21	147.6	7.8	193245	3	AC008217	AC008217 Drosophil
22	147.6	7.8	230940	3	AE003762	AE003762 Drosophil
23	144.4	7.6	1832	3	AF106964	AF106964 Strongylo
24	96	5.1	5315	3	DROE74B	M37083 Drosophila
25	96	5.1	6479	3	DROE74A	M37082 Drosophila
c 26	96	5.1	51405	2	AC019512	AC019512 Drosophil
c 27	96	5.1	174643	3	AC009371	AC009371 Drosophil
28	94.4	5.0	4800	3	DM74E	X15087 Drosophila
29	87.2	4.6	148462	2	AC095263	AC095263 Rattus no
30	83.2	4.4	2266	6	AR073028	AR073028 Sequence
31	83.2	4.4	2266	9	HUMELK1A	M25269 Homo sapien
32	83.2	4.4	2695	9	AB016193	AB016193 Homo sapi
33	81.6	4.3	848	5	AF075706	AF075706 Gallus ga
34	81.4	4.3	125020	9	AF429315	AF429315 Homo sapi
35	81.2	4.3	742	9	AF000672	AF000672 Homo sapi
36	80.8	4.3	1436	5	DRE249590	AJ249590 Danio rer
37	80.4	4.2	660	10	MMELK1P	Z36939 M.musculus
38	80.4	4.2	2286	6	AX305989	AX305989 Sequence
39	80.4	4.2	2286	10	MMELK1GEN	X87257 M.musculus
40	79.6	4.2	2508	5	AF075708	AF075708 Gallus ga
41	79.2	4.2	4200	5	DRU84615	U84615 Danio rerio
c 42	78.8	4.2	742	9	HSA339670	AJ339670 Homo sapi
43	78.8	4.2	1403	9	HSDNAFEV3	Y08979 H.sapiens F
44	78.8	4.2	1901	9	HSRNAFEV	Y08976 H.sapiens m
c 45	78.8	4.2	169741	9	AC097468	AC097468 Homo sapi

ALIGNMENTS

RESULT 1
AX048148
LOCUS AX048148 1894 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 1 from Patent WO0070092.
ACCESSION AX048148
VERSION AX048148.1 GI:11876967
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1894)
AUTHORS Kaufmann,J. and Xin,H.
TITLE Expression of ets-domain proteins in cancer
JOURNAL Patent: WO 0070092-A 1 23-NOV-2000;
Chiron Corporation (US)
FEATURES Location/Qualifiers
source 1. .1894
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 416. .1423
/note="Human JKETS."
/codon_start=1

```
                  /protein_id="CAC19002.1"
                  /db_xref="GI:11876968"
                  /translation="MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWS
                  PSPPATPEQGLSAFYLSYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPA
                  GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLL
                  WTEHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKER
                  TSPGAIHYCASTSEESWTDSEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEK
                  GIFKIEDSAQVARLWGIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP
                  I"
BASE COUNT      368 a    653 c    571 g    302 t
ORIGIN


  Query Match           98.7%;  Score 1870;  DB 6;  Length 1894;
  Best Local Similarity 99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 180

Qy    181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 240

Qy    241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    301 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    661 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1321 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1560

Qy   1561 gcaaccaactgcccccaggggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCATGGGGCAGGG 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1801 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 CCCGGGAATGGATAATAAAGATACTAGAGAACTG 1894
```

```
RESULT   2
AF071538
LOCUS       AF071538                1894 bp    mRNA    linear   PRI 10-JAN-2000
DEFINITION  Homo sapiens Ets transcription factor PDEF (PDEF) mRNA, complete
            cds.
ACCESSION   AF071538
VERSION     AF071538.1  GI:4007417
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1894)
  AUTHORS   Oettgen,P., Finger,E., Sun,Z., Akbarali,Y., Thamrongsak,U.,
            Boltax,J., Grall,F., Dube,A., Weiss,A., Brown,L., Quinn,G., Kas,K.,
            Endress,G., Kunsch,C. and Libermann,T.A.
  TITLE     PDEF, a novel prostate epithelium-specific ets transcription
            factor, interacts with the androgen receptor and activates
            prostate-specific antigen gene expression
  JOURNAL   J. Biol. Chem. 275 (2), 1216-1225 (2000)
  MEDLINE   20092894
REFERENCE   2  (bases 1 to 1894)
  AUTHORS   Libermann,T.A., Oettgen,P., Kunsch,C. and Finger,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUN-1998) Medicine, Beth Israel Deaconess Medical
            Center, 330 Brookline Ave., Boston, MA 02215, USA
FEATURES             Location/Qualifiers
     source          1. .1894
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.3"
                     /tissue_type="prostate"
     gene            1. .1894
                     /gene="PDEF"
     CDS             416. .1423
                     /gene="PDEF"
                     /note="prostate epithelium-specific"
                     /codon_start=1
                     /product="Ets transcription factor PDEF"
                     /protein_id="AAC95296.1"
                     /db_xref="GI:4007418"
                     /translation="MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWS
                     PSPPATPEQGLSAFYLSYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPA
                     GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLL
                     WTEHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKER
                     TSPGAIHYCASTSEESWTDSEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEK
                     GIFKIEDSAQVARLWGIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP
                     I"
BASE COUNT      368 a    653 c    571 g    302 t
ORIGIN
```

```
  Query Match             98.7%;  Score 1870;  DB 9;  Length 1894;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 180

Qy    181 gccaccagccacccccttgagggtggccaggccccagtggccaacctgagtgctgcctct 240
          |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 240

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    301 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCGACAC 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    661 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgccccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
```

```
Db   1321 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
Db   1381 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
Db   1441 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
Db   1501 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
Db   1561 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCATGGGGCAGGG 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
Db   1621 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
Db   1681 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
Db   1741 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
Db   1801 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
Db   1861 CCCGGGAATGGATAATAAAGATACTAGAGAACTG 1894
```

```
RESULT   3
AX175516
LOCUS       AX175516                1918 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 1 from Patent WO0142472.
ACCESSION   AX175516
VERSION     AX175516.1  GI:14598855
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1918)
  AUTHORS   Harkins,R., Lin,R.J., Luke,M., Monteclaro,F., Parkes,D., Parry,G.,
            Steinbrecher,R., van Heuit,P.T. and Xuan,J.A.
  TITLE     Dna encoding a novel prost-ets polypeptide
  JOURNAL   Patent: WO 0142472-A 1 14-JUN-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .1918
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             422. .1429
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC43877.1"
                     /db_xref="GI:14598856"
                     /translation="MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWS
                     PSPPATPEQGLSAFYLSYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPA
                     GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLL
                     WTEHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKER
                     TSPGAIHYCASTSEESWTDSEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEK
                     GIFKIEDSAQVARLWGIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP
                     I"
```

```
BASE COUNT      385 a    656 c    574 g    303 t
ORIGIN

  Query Match           98.6%;  Score 1868.4;  DB 6;  Length 1918;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1878;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
Db       7 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 66

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
Db      67 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 126

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
Db     127 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 186

Qy     181 gccaccagcccacccctgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
Db     187 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 246

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
Db     247 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 306

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
Db     307 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 366

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
Db     367 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 426

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
Db     427 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 486

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
Db     487 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 546

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
Db     547 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 606

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
Db     607 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 666

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
Db     667 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 726

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
Db     727 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 786

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
Db     787 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 846

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
Db     847 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 906

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
Db     907 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 966

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
```

```
Db    967 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1026

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
Db   1027 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1086

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
Db   1087 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1146

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
Db   1147 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1206

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
Db   1207 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1266

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
Db   1267 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1326

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
Db   1327 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1386

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
Db   1387 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1446

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
Db   1447 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1506

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
Db   1507 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1566

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
Db   1567 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGG 1626

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
Db   1627 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1686

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
Db   1687 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1746

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
Db   1747 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1806

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
Db   1807 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1866

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
Db   1867 CCCGGGAATGGATAATAAAGATACTAGAGAACTG 1900
```

```
RESULT   4
AX048159
LOCUS       AX048159               3317 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 12 from Patent WO0070092.
ACCESSION   AX048159
VERSION     AX048159.1  GI:11876975
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3317)
  AUTHORS   Kaufmann,J. and Xin,H.
  TITLE     Expression of ets-domain proteins in cancer
  JOURNAL   Patent: WO 0070092-A 12 23-NOV-2000;
            Chiron Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .3317
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="GSEF-encoding sequence with promoter (Figs. 2A-2B)"
     promoter        1. .1380
                     /note="Promoter"
     CDS             1815. .2822
                     /note="GSEF coding sequence"
                     /codon_start=1
                     /protein_id="CAC19004.1"
                     /db_xref="GI:11876976"
                     /translation="MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWS
                     PSPPATPEQGLSAFYLSYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPA
                     GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLL
                     WTEHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKER
                     TSPGAIHYCASTSEESWTDSEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEK
                     GIFKIEDSAQVARLWGIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP
                     I"
BASE COUNT      710 a   1026 c    970 g    611 t
ORIGIN
```

```
  Query Match             98.5%;  Score 1865.2;  DB 6;  Length 3317;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches 1876;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
Db   1400 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 1459

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
Db   1460 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 1519

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
Db   1520 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 1579

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
Db   1580 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 1639

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
Db   1640 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 1699

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
Db   1700 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 1759

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
Db   1760 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 1819

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
Db   1820 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 1879

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
Db   1880 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 1939

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
Db   1940 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 1999
```

```
Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2000 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 2059

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db   2060 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 2119

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2120 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 2179

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2180 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 2239

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2240 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 2299

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2300 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 2359

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   2360 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGAGATGTGCT 2419

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2420 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 2479

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2480 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 2539

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 2599

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2600 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 2659

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2660 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 2719

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   2720 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 2779

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 2839

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 2899

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   2900 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 2959

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   2960 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGG 3019

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3020 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 3079

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3080 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 3139

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3140 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 3199

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   3200 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 3259

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||| ||||||||||||||||
Db   3260 CCCGGGAATGGATAATAGAGATACTAGAGAACTG 3293
```

RESULT 5
BC021299
LOCUS BC021299 1942 bp mRNA linear PRI 22-JAN-2002
DEFINITION Homo sapiens, prostate epithelium-specific Ets transcription factor, clone MGC:29592 IMAGE:4546860, mRNA, complete cds.
ACCESSION BC021299
VERSION BC021299.1 GI:18204265
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1942)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R., Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J., Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 39 Row: d Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6912579.

FEATURES Location/Qualifiers
source 1. .1942
/organism="Homo sapiens"
/db_xref="LocusID:25803"
/db_xref="taxon:9606"
/clone="MGC:29592 IMAGE:4546860"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_15"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
CDS 421. .1428
/codon_start=1
/product="prostate epithelium-specific Ets transcription

```
                     factor"
                     /protein_id="AAH21299.1"
                     /db_xref="GI:18204266"
                     /translation="MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWS
                     PSPPATPEQGLSAFYLSYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPA
                     GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLL
                     WTEHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKER
                     TSPGAIHYCASTSEESWTDSEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEK
                     GIFKIEDSAQVARLWGIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP
                     I"
BASE COUNT      411 a    656 c    575 g    300 t
ORIGIN


  Query Match           98.4%;  Score 1864.4;  DB 9;  Length 1942;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1874;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       5 gacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtc 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10 GACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCCAGTC 69

Qy      65 ctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccag 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      70 CTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTGCCAG 129

Qy     125 caggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagaccgcca 184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     130 CAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACCTCAG 189

Qy     185 ccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgcca 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     190 AGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCTGCCA 249

Qy     245 ccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgccagt 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     250 CCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGCCAGT 309

Qy     305 ggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacagctcc 364
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     310 GGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAGCTCC 369

Qy     365 ctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgggcagc 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     370 CTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGGCAGC 429

Qy     425 gccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacacggtg 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     430 GCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACACGGTG 489

Qy     485 tcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggactgg 544
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     490 TCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGACTGG 549

Qy     545 agtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcctacttt 604
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     550 AGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTACTTT 609

Qy     605 gacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcagtcgg 664
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     610 GACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAGTCGG 669

Qy     665 gaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagcgggc 724
           |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db     670 GAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGCGGGC 729

Qy     725 accctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggtgcag 784
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     730 AGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGTGCAG 789

Qy     785 tccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaacatc 844
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     790 TCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAACATC 849

Qy     845 accgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggacagag 904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     850 ACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGACAGAG 909

Qy     905 caccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaaggagctg 964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     910 CACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGAGCTG 969

Qy     965 tgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgctgcac 1024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     970 TGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCTGCAC 1029

Qy    1025 gcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctggggcg 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1030 GCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGGGGCG 1089

Qy    1085 attcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactcatca 1144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1090 ATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCA 1149

Qy    1145 tgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccac 1204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1150 TGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCAC 1209

Qy    1205 agctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggac 1264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1210 AGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGAGGAC 1269

Qy    1265 tcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgac 1324
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    1270 TCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTACGAC 1329

Qy    1325 aagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccagacatc 1384
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db    1330 AAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGACATC 1389

Qy    1385 tcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaa 1444
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1390 TCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCCTGAA 1449

Qy    1445 acccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatggggg 1504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1450 ACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGG 1509

Qy    1505 aaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggggcaa 1564
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    1510 AAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAA 1569

Qy    1565 ccaactgccccaggggggatatgggtcctctggggccttcgggaccatggggcaggggtgc 1624
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1570 CCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGGGTGC 1629

Qy    1625 ttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccccaac 1684
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1630 TTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAAC 1689

Qy    1685 acctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaa 1744
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 ACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAA 1749

Qy    1745 aggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgcaccac 1804
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1750 AGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCACCAC 1809

Qy    1805 acctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgccccca 1864
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1810 ACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCCCCCG 1869

Qy    1865 ggaatggataataaagatactagagaactg 1894
           ||||||||||||||||||||||||||||||
Db    1870 GGAATGGATAATAAAGATACTAGAGAACTG 1899
```

```
RESULT   6
AB031549
LOCUS       AB031549                1898 bp    mRNA    linear   PRI 20-JAN-2000
DEFINITION  Homo sapiens PSE mRNA for prostate ets, complete cds.
ACCESSION   AB031549
VERSION     AB031549.1  GI:6721497
KEYWORDS    prostate ets.
SOURCE      Homo sapiens adult male prostate epithelium prostate cancer
            cell_line:PC-3 cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Yamada,N., Tamai,Y., Miyamoto,H. and Nozaki,M.
  TITLE     Cloning and expression of the mouse Pse gene encoding a novel Ets
            family member
  JOURNAL   Gene 241 (2), 267-274 (2000)
  MEDLINE   20137504
REFERENCE   2  (bases 1 to 1898)
  AUTHORS   Nozaki,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-AUG-1999) Masami Nozaki, Osaka University, Institute
            for Microbial Diseases; 3-1, Yamadaoka, Suita, Osaka 565-0871,
            Japan (E-mail:mnozaki@biken.osaka-u.ac.jp, Tel:81-6-879-8338,
            Fax:81-6-879-8339)
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p21.2-21.3"
                     /sex="male"
                     /cell_line="PC-3"
                     /cell_type="prostate cancer"
                     /tissue_type="prostate epithelium"
                     /dev_stage="adult"
                     /note="vector:pUC18"
     gene            406. .1413
                     /gene="PSE"
     CDS             406. .1413
                     /gene="PSE"
                     /codon_start=1
                     /product="prostate ets"
                     /protein_id="BAA89543.1"
                     /db_xref="GI:6721498"
                     /translation="MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWS
                     PSPPATPEQGLSAFYLSYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPA
                     GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLL
                     WTEHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKER
                     TSPGAIHYCASTSEESWTDSEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEK
                     GIFKIEDSAQVARLWGIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP
                     I"
BASE COUNT      380 a    651 c    567 g    300 t
ORIGIN
```

```
  Query Match             95.7%;  Score 1811.8;  DB 9;  Length 1898;
  Best Local Similarity   98.9%;  Pred. No. 1.6e-293;
  Matches 1866;  Conservative    0;  Mismatches   17;  Indels    4;  Gaps    4;

Qy       8 ttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtcctc 67
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCCAGTCCTC 60

Qy      68 caagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccagcag 127
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTGCCAGCAG 120

Qy     128 gcagccctgggctgggggtaggggactccctacaggcacgcagccctgagaccgccacca 187
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     121 GCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACCTCAGAGG 180

Qy     188 gccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgccacca 247
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     181 GCCACCCCTTGAGGGTGG-CAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCTGCCACCA 239

Qy     248 gccctgctggcccctggttccgctggcccccagatgcctggctgagacacgccagtggc 307
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     240 GCCCTGCT-GCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGCCAGT-GC 297

Qy     308 ctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacagctccctg 367
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     298 CTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAGCTCCCTG 357

Qy     368 ggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatgggcagcgcc 427
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     358 GGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGGCAGCGCC 417

Qy     428 agcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacacggtgtcg 487
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 AGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACACGGTGTCG 477

Qy     488 cggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggactggagt 547
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 CGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGACTGGAGT 537

Qy     548 cccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcctactttgac 607
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     538 CCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTACTTTGAC 597

Qy     608 atgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcagtcgggag 667
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 ATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAGTCGGGAG 657

Qy     668 gagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagcgggcacc 727
           |||||||||||  |||||||||||||||||||||||||||||||||||||||||||| |
Db     658 GAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGCGGGCAGC 717

Qy     728 ctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggtgcagtcc 787
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 CTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGTGCAGTCC 777

Qy     788 atggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaacatcacc 847
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 ATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAACATCACC 837

Qy     848 gcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggacagagcac 907
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 GCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGACAGAGCAC 897

Qy     908 caataccggctgccccccatgggcaaggccttccaggagctggcgggcaaggagctgtgc 967
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 CAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGAGCTGTGC 957

Qy     968 gccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcc 1027
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 GCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCTGCACGCC 1017

Qy    1028 cacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctggggcgatt 1087
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 CACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGGGGCGATT 1077

Qy    1088 cactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactcatcatgc 1147
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1078 CACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCATGC 1137

Qy    1148 tccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagc 1207
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 TCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCACAGC 1197

Qy    1208 tatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactca 1267
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 TATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGAGGACTCA 1257

Qy    1268 gcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactgcgacaag 1327
```

```
Db    1258 GCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTACGACAAG 1317

Qy    1328 ctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccagacatctcc 1387
           ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    1318 CTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGACATCTCC 1377

Qy    1388 cagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcctgaaacc 1447
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1378 CAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCCTGAAACC 1437

Qy    1448 cgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatgggggaaa 1507
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1438 CGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATGGGGGAAA 1497

Qy    1508 acgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggggcaacca 1567
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 ACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGGGCAACCA 1557

Qy    1568 actgccccaggggatatgggtcctctggggccttcgggaccatggggcaggggtgcttc 1627
           |||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||
Db    1558 ACTGCCCCAGGGGGATATGGGTCCTCTGGGG-CTTCGGGACCCTGGGGCAGGGGTGCTTC 1616

Qy    1628 ctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccccaacacc 1687
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1617 CTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCCCAACACC 1676

Qy    1688 tgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcgacaaagg 1747
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1677 TGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCGACAAAGG 1736

Qy    1748 ccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgcaccacacc 1807
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1737 CCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCACCACACC 1796

Qy    1808 tggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcccccggga 1867
           |||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||
Db    1797 TGGCATGGTGCAGGGAGACATCTGCCCCCCTGAGTTGGGCAGCCAGGAGTGCCCCCGGGA 1856

Qy    1868 atggataataaagatactagagaactg 1894
           ||||||||||||||||||||||||||||
Db    1857 ATGGATAATAAAGATACTAGAGAACTG 1883
```

```
RESULT   7
BC012648
LOCUS       BC012648                1856 bp    mRNA    linear   ROD 20-AUG-2001
DEFINITION  Mus musculus, Similar to prostate specific ets transcription
            factor, clone MGC:13763 IMAGE:4218741, mRNA, complete cds.
ACCESSION   BC012648
VERSION     BC012648.1  GI:15215054
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1856)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
            Greene, Mark Ketteman and Anuradha Madan

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 18 Row: p Column: 20
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 4760536.
FEATURES             Location/Qualifiers
     source          1. .1856
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="MGC:13763 IMAGE:4218741"
                     /tissue_type="Colon, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             427. .1404
                     /codon_start=1
                     /product="Similar to prostate specific ets transcription
                     factor"
                     /protein_id="AAH12648.1"
                     /db_xref="GI:15215055"
                     /translation="MGSASPGLSNVSPGCLLLFPDVAPRTGTEKAASGAMGPEKQEWS
                     PSPPATPEQGLSAFYLSYFNMYPDDSSWVAKVPEARAGEDHPEEPEQCPVIDSQASGS
                     TLDEHSLEQVQSMVVGEVLKDIETACKLLNITADPGDWSPGNVQKWLLWTEHQYRLPP
                     AGKAFQELGGKELCAMSEEQFRQRAPLGGDVLHAHLDIWKSAAWMKERTSPGTLHYCA
                     STSEEGWTDGEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQ
                     VARLWGVRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPV"
BASE COUNT      421 a    575 c    520 g    340 t
ORIGIN
```

```
  Query Match             40.0%;  Score 757;  DB 10;  Length 1856;
  Best Local Similarity   73.2%;  Pred. No. 3.7e-117;
  Matches 1087;  Conservative    0;  Mismatches  345;  Indels   53;  Gaps    7;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||| |  ||||| |  . | |  | |    |       |||||  | |||||  | || |||
Db     31 GTCCGCATTCCTGCAGTCCCTCTGCCTTCCGTCCATCTGTCCGCTCTGCCTCCCACACCT 90

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||  | |||  ||      | || ||    |   ||| ||| ||||||||||| || ||||
Db     91 AGCACCCCAGCCC--GCTGCTGCCCCGGTGAGAACCCCCAGCTTGGGCCTTGTCATGGTG 148

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||  |||||| |||    |  ||||| || |||| || |  |||| |||||| ||
Db    149 CCAGCAGGTGGCCCTGAGCTTCTGACAGGGGCCTGCCTATAGACCTGCAGGCCTGAGGCC 208

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
              ||| | | |  | | || |     | |||  |||||| |||| || ||  |||||
Db    209 ----TCAGACTCACACTCAAGGGGCAAGAGGCCCTGGTGGCCCACCTAAGAGCCACCTCT 264

Qy    241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
          | | ||||||||||||||||||      ||| |           | |  ||     |
Db    265 GTCCCCAGCCCTGCTGGCCCCACTGATGTCTGAC----------TGAGACCCAGCAGTGA 314

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |  ||  ||      ||    |||| ||| || |||||  ||||||  |||| |  |||
Db    315 CCCTGAGCTGCCTGCCCACTGCCTCCTCCTGGTCCCTGAGGTTGGCTCTGCCGAGGACGA 374

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          |||    | || |||||  ||||||  ||  |||||  | ||||   ||  |||   |||||||
Db    375 CTCTTCTGACAGCAGGCGGCTAACGGAAGCAGCCCCAAGCCTCCACCGCA---GCATGGG 431

Qy    421 cagcgccagcccgggtctgagcagcgtatccccagccacctcctgctgccccccgacac 480
          |||  |||||||| || |||||||  ||| ||||||  |   |||  ||  |||  ||| ||
Db    432 CAGTGCCAGCCCAGGCCTGAGCAACGTGTCCCCCGGTTGCCTGCTACTGTTCCCAGATGT 491

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgggga 540
          ||   | ||  |||||   |||||||||  ||| | ||  |||  ||||  |  ||||  | |||
Db    492 GGCACCACGAACAGGGACGGAGAAGGCAGCATCAGGAGCAATGGGCCCTGAGAAGCAGGA 551

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
```

```
Db    552 ATGGAGTCCTAGTCCACCCGCCACCCCTGAGCAGGGCCTGTCTGCTTTCTACCTCTCTTA 611

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
Db    612 CTTTAACA---TGTATCCCGACGATAGCAGCTGGGTCGCCAAAGTCCCCGAGGCCCGTGC 668

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
Db    669 CGGGGAGGACCACCCGGAGGAGCCCGAGCAGTGTCCCGTCATTGACAGCCAGGCCTCTGG 728

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
Db    729 GAGC-------------------------ACGTTGGATGAGCACTCGCTAGAGCAGGT 761

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
Db    762 GCAATCGATGGTTGTGGGCGAGGTCCTGAAAGATATTGAGACGGCCTGCAAGCTTCTGAA 821

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
Db    822 CATCACAGCAGACCCTGGGGACTGGAGCCCTGGTAACGTGCAGAAGTGGCTTTTATGGAC 881

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
Db    882 AGAACACCAGTACCGGCTGCCTCCAGCAGGCAAGGCCTTCCAGGAGCTGGGCGGTAAGGA 941

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
Db    942 GCTGTGCGCCATGTCCGAGGAACAGTTCCGTCAGCGTGCACCCTTGGGTGGGGATGTACT 1001

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
Db   1002 GCATGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAGGAGAGGACCTCGCCTGG 1061

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
Db   1062 GACCCTTCACTACTGCGCCTCCACCAGCGAGGAGGGCTGGACGGATGGTGAGGTGGACTC 1121

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
Db   1122 GTCGTGCTCCGGGCAGCCCATTCACCTGTGGCAGTTCCTGAAAGAACTGCTGCTCAAGCC 1181

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
Db   1182 CCACAGCTATGGCCGCTTCATCCGCTGGCTCAACAAGGAGAAAGGCATCTTCAAAATTGA 1241

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
Db   1242 GGACTCAGCACAGGTGGCCCGACTGTGGGGTGTGCGCAAGAACCGGCCAGCCATGAACTA 1301

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
Db   1302 TGATAAACTAAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATTCGTAAACCCGA 1361

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcc----tggcccag 1436
Db   1362 CATCTCTCAGCGCCTTGTCTACCAATTTGTGCATCCAGTCTGAGAGCCACAGAGACCAGA 1421

Qy   1437 ggcctgaaaacccgccctcaggggcctctctcctgcctgccctgcc 1481
Db   1422 GGCCTACAACCTGCCCCAGGCAGCCACTCTCTGGTTGGCCTGGTC 1466
```

```
RESULT   8
AB019436
LOCUS       AB019436               1704 bp    mRNA    linear   ROD 20-JAN-2000
DEFINITION  Mus musculus mRNA for Prostate specific transcription factor ets,
            complete cds.
ACCESSION   AB019436
VERSION     AB019436.1  GI:4760536
KEYWORDS    Prostate specific transcription factor ets; mPse.
SOURCE      Mus musculus (strain:C57BL/6) Adult male Colon Epithelial cDNA to
            mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (sites)
  AUTHORS   Yamada,N., Tamai,Y., Miyamoto,H. and Nozaki,M.
  TITLE     Cloning and expression of the mouse Pse gene encoding a novel Ets
            family member
  JOURNAL   Gene 241 (2), 267-274 (2000)
  MEDLINE   20137504
REFERENCE   2  (bases 1 to 1704)
  AUTHORS   Nozaki,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1998) Masami Nozaki, Osaka University, Institute
            for Microbial Diseases; 3-1, Yamadaoka, Suita, Osaka 565-0871,
            Japan (E-mail:mnozaki@biken.osaka-u.ac.jp, Tel:81-6-879-8338,
            Fax:81-6-879-8339)
FEATURES             Location/Qualifiers
     source          1. .1704
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /sex="male"
                     /cell_type="Epithelial"
                     /tissue_type="Colon"
                     /dev_stage="Adult"
     gene            362. .1339
                     /gene="mPse"
     CDS             362. .1339
                     /gene="mPse"
                     /codon_start=1
                     /product="Prostate specific transcription factor ets"
                     /protein_id="BAA77329.1"
                     /db_xref="GI:4760537"
                     /translation="MGSASPGLSNVSPGCLLLFPDVAPRTGTEKAASGAMGPEKQEWS
                     PSPPATPEQGLSAFYLSYFNMYPDDSSWVAKVPEARAGEDHPEEPEQCPVIDSQASGS
                     TLDEHSLEQVQSMVVGEVLKDIETACKLLNITADPGDWSPGNVQKWLLWTEHQYRLPP
                     AGKAFQELGGKELCAMSEEQFRQRAPLGGDVLHAHLDIWKSAAWMKERTSPGTLHYCA
                     STSEEGWTDGEVDSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQ
                     VARLWGVRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPV"
     polyA_site      1704
                     /note="10 a nucleotides"
BASE COUNT      369 a    527 c    491 g    317 t
ORIGIN
```

```
  Query Match             39.0%;  Score 738.4;  DB 10;  Length 1704;
  Best Local Similarity   73.3%;  Pred. No. 4.8e-114;
  Matches 1061;  Conservative    0;  Mismatches  336;  Indels   51;  Gaps    7;

Qy      38 tgtccacactgccccacagacccagtcctccaagcctgctgccagctccctgcaagcccc 97
Db       1 TGTCCGCTCTGCCTCCCACACCTAGCACCCCAGCCC--GCTGCTGCCCCCGGTGAGAACCC 58

Qy      98 tcaggttgggccttgccacggtgccagcaggcagccctgggctgggggtaggggactccc 157
Db      59 CCAGCTTGGGCCTTGTCATGGTGCCAGCAGGTGGCCCTGAGCTTCTGACAGGGGCCTGCC 118

Qy     158 tacaggcacgcagccctgagaccgccaccagccacccccttgagggtggccaggcccccag 217
Db     119 TATAGACCTGCAGGCCTGAGGCC----TCAGACTCACACTCAAGGGGCAAGAGGCCCTGG 174

Qy     218 tggccaacctgagtgctgcctctgccaccagccctgctggcccctggttccgctggcccc 277
Db     175 TGGCCCACCTAAGAGCCACCTCTGTCCCCAGCCCTGCTGCCCCACTGAT------GTCTG 228

Qy     278 ccagatgcctggctgagacacgccagtggcctcagctgcccacacctcttcccggccccct 337
Db     229 ACTGAGACCCAGCAGTGACCC-----TGAGCTGCCTGCCCACTGCCTCCTCCTGGTCCCT 283

Qy     338 ggagttggcactgcagcagacagctccctgggcaccaggcagctaacagacacagccgcc 397
Db     284 GAGGTTGGCTCTGCCGAGGACGGACGACTCTTCTGAAGCAGGCGGCTAACGGAAGCAGCC 343
```

```
Qy    398 agcccaaacagcagcggcatgggcagcgccagcccgggtctgagcagcgtatcccccagc 457
Db    344 CCAAGCCTCCACCGCAGCATGGGCAGTGCCAGCCCAGGCCTGAGCAACGTGTCCCCCGGT 403

Qy    458 cacctcctgctgccccccgacacggtgtcgcggacaggcttggagaaggcggcagcgggg 517
Db    404 TGCCTGCTACTGTTCCCAGATGTGGCACCACGAACAGGGACGGAGAAGGCAGCATCAGGA 463

Qy    518 gcagtgggtctcgagagacgggactggagtcccagtccacccgccacgcccgagcagggc 577
Db    464 GCAATGGGCCCTGAGAAGCAGGAATGGAGTCCTAGTCCACCCGCCACCCCTGAGCAGGGC 523

Qy    578 ctgtccgccttcgacctctcctactttgacatgctgtaccctgaggacagcagctgggca 637
Db    524 CTGTCTGCTTTCTACCTCTCTTACTTTAACA---TGTATCCCGACGATAGCAGCTGGGTC 580

Qy    638 gccaaggcccctggggccagcagtcgggaggagccacctgaccagcctgagcagtgcccg 697
Db    581 GCCAAAGTCCCCGAGGCCCGTGCCGGGGAGGACCACCCGGAGGAGCCCGAGCAGTGTCCC 640

Qy    698 gtcattgacagccaagccccagcgggcaccctggacttggtgcccggcgggctgaccttg 757
Db    641 GTCATTGACAGCCAGGCCTCTGGGAGC----------------------------ACGTTG 673

Qy    758 gaggagcactcgctggagcaggtgcagtccatggtggtgggcgaagtgctcaaggacatc 817
Db    674 GATGAGCACTCGCTAGAGCAGGTGCAATCGATGGTTGTGGGCGAGGTCCTGAAAGATATT 733

Qy    818 gagacggcctgcaagctgctcaacatcaccgcagatcccatggactggagccccagcaat 877
Db    734 GAGACGGCCTGCAAGCTTCTGAACATCACAGCAGACCCTGGGGACTGGAGCCCTGGTAAC 793

Qy    878 gtgcagaagtggctcctgtggacagagcaccaataccggctgccccccatgggcaaggcc 937
Db    794 GTGCAGAAGTGGCTTTTATGGACAGAACACCAGTACCGGCTGCCTCCAGCAGGCAAGGCC 853

Qy    938 ttccaggagctggcgggcaaggagctgtgcgccatgtcggaggagcagttccgccagcgc 997
Db    854 TTCCAGGAGCTGGGCGGTAAGGAGCTGTGCGCCATGTCCGAGGAACAGTTCCGTCAGCGT 913

Qy    998 tcgcccctgggtggggatgtgctgcacgcccacctggacatctggaagtcagcggcctgg 1057
Db    914 GCACCCTTGGGTGGGGATGTACTGCATGCCCACCTGGACATCTGGAAGTCAGCGGCCTGG 973

Qy   1058 atgaaagagcggacttcacctggggcgattcactactgtgcctcgaccagtgaggagagc 1117
Db    974 ATGAAGGAGAGGACCTCGCCTGGGACCCTTCACTACTGCGCCTCCACCAGCGAGGAGGGC 1033

Qy   1118 tggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctgtggcagttc 1177
Db   1034 TGGACGGATGGTGAGGTGGACTCGTCGTGCTCCGGGCAGCCCATTCACCTGTGGCAGTTC 1093

Qy   1178 ctcaaggagttgctactcaagccccacagctatggccgcttcattaggtggctcaacaag 1237
Db   1094 CTGAAAGAACTGCTGCTCAAGCCCCACAGCTATGGCCGCTTCATCCGCTGGCTCAACAAG 1153

Qy   1238 gagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgc 1297
Db   1154 GAGAAAGGCATCTTCAAAATTGAGGACTCAGCACAGGTGGCCCGACTGTGGGGTGTGCGC 1213

Qy   1298 aagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaag 1357
Db   1214 AAGAACCGGCCAGCCATGAACTATGATAAACTAAGCCGCTCCATCCGCCAGTATTACAAG 1273

Qy   1358 aagggcatcatccggaagccagacatctcccagcgcctcgtctaccagttcgtgcacccc 1417
Db   1274 AAGGGCATCATTCGTAAACCCGACATCTCTCAGCGCCTTGTCTACCAATTTGTGCATCCA 1333

Qy   1418 atctgagtgcc----tggcccagggcctgaaacccgccctcaggggcctctctcctgcct 1473
Db   1334 GTCTGAGAGCCACAGAGACCAGAGGCCTACAACCTGCCCCAGGCAGCCACTCTCTGGTTG 1393
```

```
Qy   1474 gccctgcc 1481
          ||| | |
Db   1394 GCCTGGTC 1401
```

```
RESULT   9
AL157372/c
LOCUS       AL157372              150660 bp    DNA     linear   PRI 29-NOV-2000
DEFINITION  Human DNA sequence from clone RP11-375E1 on chromosome 6, complete
            sequence.
ACCESSION   AL157372
VERSION     AL157372.18  GI:11493236
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 150660)
  AUTHORS   Skuce,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 30, 2000 this sequence version replaced gi:11342772.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows.  An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated repeat sequence elements.  Where the sequence is
            ambiguous, there is an annotation using the 'unsure' feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-375E1 is from the library RPCI-11.2 constructed at the Roswell
            Park Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-375E1 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-375E1 is at 1 in this sequence. The
            true left end of clone RP3-391O22 is at 150561 in this sequence.
            The true right end of clone RP11-302L10 is at 113044 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1. .150660
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-375E1"
                     /clone_lib="RPCI-11.2"
BASE COUNT    35470 a  38887 c  38740 g  37563 t
ORIGIN
```

```
  Query Match           33.9%;  Score 643;  DB 9;  Length 150660;
  Best Local Similarity   99.2%;  Pred. No. 1.5e-98;
  Matches  646;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   1244 ggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaac 1303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123739 GGCATCTTCAAAATTGAGGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAAC 123680
```

```
Qy   1304 cgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaagaagggc 1363
          ||||||||||||||||| |||||||||||||||||||||||||||| ||||||||||||
Db 123679 CGTCCCGCCATGAACTACGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGC 123620

Qy   1364 atcatccggaagccagacatctcccagcgcctcgtctaccagttcgtgcaccccatctga 1423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123619 ATCATCCGGAAGCCAGACATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGA 123560

Qy   1424 gtgcctggcccagggcctgaaacccgccctcaggggcctctctcctgcctgccctgcctc 1483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123559 GTGCCTGGCCCAGGGCCTGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTC 123500

Qy   1484 agccaggccctgagatgggggaaaacgggcagtgtgctctgctgctctgaccttccagag 1543
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 123499 AGCCAGGCCCTGAGATGGGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAG 123440

Qy   1544 cccaaggtcagggagggggcaaccaactgccccaggggatatgggtcctctggggccttc 1603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123439 CCCAAGGTCAGGGAGGGGCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTC 123380

Qy   1604 gggaccatggggcaggggtgcttcctcctcaggcccagctgctcccctggaggacagagg 1663
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123379 GGGACCCTGGGGCAGGGGTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGG 123320

Qy   1664 gagacagggctgctccccaacacctgcctctgaccccagcatttccagagcagagcctac 1723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123319 GAGACAGGGCTGCTCCCCAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTAC 123260

Qy   1724 agaagggcagtgactcgacaaaggccacaggcagtccaggcctctctctgctccatcccc 1783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123259 AGAAGGGCAGTGACTCGACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCC 123200

Qy   1784 ctgcctcccattctgcaccacacctggcatggtgcagggagacatctgcacccctcagtt 1843
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 123199 CTGCCTCCCATTCTGCACCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTT 123140

Qy   1844 gggcagccaggagtgccccccgggaatggataataaagatactagagaactg 1894
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 123139 GGGCAGCCAGGAGTGCCCCCGGGAATGGATAATAAAGATACTAGAGAACTG 123089


RESULT 10
AX048155
LOCUS       AX048155               1751 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 8 from Patent WO0070092.
ACCESSION   AX048155
VERSION     AX048155.1  GI:11876971
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1751)
  AUTHORS   Kaufmann,J. and Xin,H.
  TITLE     Expression of ets-domain proteins in cancer
  JOURNAL   Patent: WO 0070092-A 8 23-NOV-2000;
            Chiron Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .1751
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     promoter        1. .1381
                     /note="JKETS promoter"
     TATA_signal     1356. .1360
                     /note="TATA box."
BASE COUNT      369 a    521 c    497 g    364 t
ORIGIN


  Query Match             17.4%;  Score 329.8;  DB 6;  Length 1751;
  Best Local Similarity   97.7%;  Pred. No. 9.8e-46;
  Matches  345;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1400 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 1459

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1460 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 1519

Qy    121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1520 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 1579

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |    ||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1580 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGT-GCCAACCTGAGTGCTGCCTCT 1638

Qy    241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1639 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 1698

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcag 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1699 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAG 1751


RESULT 11
AX106617
LOCUS       AX106617               278 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 398 from Patent WO0125272.
ACCESSION   AX106617
VERSION     AX106617.1  GI:13922282
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 278)
  AUTHORS   Xu,J., Skeiky,Y.A., Reed,S.G. and Cheever,M.A.
  TITLE     Compositions and methods for therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0125272-A 398 12-APR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1. .278
                     /note="n = A,T,C or G"
BASE COUNT       56 a     85 c     87 g     49 t      1 others
ORIGIN


  Query Match             13.9%;  Score 263;  DB 6;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-34;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 AGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCTGCACGCCCACCTGGACATCT 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 GGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGGGGCGATTCACTACTGTGCCT 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 CGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCATGCTCCGGGCAGCCCA 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCACAGCTATGGCCGCTTCA 254
```

```
Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 TTANGTGGCTCAACAAGGAGAAGG 278


RESULT 12
AX140908
LOCUS       AX140908                 278 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 398 from Patent WO0134802.
ACCESSION   AX140908
VERSION     AX140908.1  GI:14281005
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 278)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Retter,M.W., Stolk,J.A., Day,C.H.,
            Skeiky,Y.A. and Wang,A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0134802-A 398 17-MAY-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       56 a     85 c     87 g     49 t      1 others
ORIGIN


  Query Match             13.9%;  Score 263;  DB 6;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-34;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 AGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCTGCACGCCCACCTGGACATCT 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 GGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGGGGCGATTCACTACTGTGCCT 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 CGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCATGCTCCGGGCAGCCCA 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCACAGCTATGGCCGCTTCA 254

Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 TTANGTGGCTCAACAAGGAGAAGG 278


RESULT 13
AX200768
LOCUS       AX200768                 278 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 398 from Patent WO0151633.
ACCESSION   AX200768
VERSION     AX200768.1  GI:15390661
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 278)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Retter,M.W.,
            Stolk,J.A., Skeiky,Y.A., Wang,A. and Meagher,M.J.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0151633-A 398 19-JUL-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       56 a     85 c     87 g     49 t      1 others
ORIGIN


  Query Match             13.9%;  Score 263;  DB 6;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-34;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 AGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCTGCACGCCCACCTGGACATCT 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 GGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGGGGCGATTCACTACTGTGCCT 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 CGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCATGCTCCGGGCAGCCCA 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCACAGCTATGGCCGCTTCA 254

Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 TTANGTGGCTCAACAAGGAGAAGG 278


RESULT 14
AX267424
LOCUS       AX267424                 278 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 398 from Patent WO0173032.
ACCESSION   AX267424
VERSION     AX267424.1  GI:16516188
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
            Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A., Hepler,W.T.
            and Henderson,R.A.
  TITLE     Compositions and methods for the therapy and diagnosis of prostate
            cancer
  JOURNAL   Patent: WO 0173032-A 398 04-OCT-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       56 a     85 c     87 g     49 t      1 others
ORIGIN


  Query Match             13.9%;  Score 263;  DB 6;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-34;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 AGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCTGCACGCCCACCTGGACATCT 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 GGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGGGGCGATTCACTACTGTGCCT 134

Qy    1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 CGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTCATCATGCTCCGGGCAGCCCA 194

Qy    1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 TCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCCCCACAGCTATGGCCGCTTCA 254

Qy    1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db     255 TTANGTGGCTCAACAAGGAGAAGG 278


RESULT  15
AC095858
LOCUS       AC095858               105772 bp    DNA     linear   HTG 20-DEC-2001
DEFINITION  Rattus norvegicus clone CH230-10E14, *** SEQUENCING IN PROGRESS
            ***, 40 unordered pieces.
ACCESSION   AC095858
VERSION     AC095858.3  GI:17943471
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 105772)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 105772)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Dec 20, 2001 this sequence version replaced gi:16901667.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GDQD
                Center clone name: CH230-10E14
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329First call to
            findPhrapList
                Consensus quality: 84212 bases at least Q40
                Consensus quality: 90046 bases at least Q30
                Consensus quality: 95305 bases at least Q20
                Estimated insert size: 90346; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 1.1x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 40 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     7182: contig of 7182 bp in length
            *     7183     7282: gap of unknown length
            *     7283    12821: contig of 5539 bp in length
            *    12822    12921: gap of unknown length
            *    12922    17580: contig of 4659 bp in length
            *    17581    17680: gap of unknown length
            *    17681    21052: contig of 3372 bp in length
            *    21053    21152: gap of unknown length
            *    21153    24730: contig of 3578 bp in length
            *    24731    24830: gap of unknown length
            *    24831    26079: contig of 1249 bp in length
            *    26080    26179: gap of unknown length
            *    26180    30636: contig of 4457 bp in length
            *    30637    30736: gap of unknown length
            *    30737    34528: contig of 3792 bp in length
            *    34529    34628: gap of unknown length
            *    34629    37889: contig of 3261 bp in length
            *    37890    37989: gap of unknown length
            *    37990    41096: contig of 3107 bp in length
            *    41097    41196: gap of unknown length
            *    41197    44133: contig of 2937 bp in length
            *    44134    44233: gap of unknown length
            *    44234    47711: contig of 3478 bp in length
            *    47712    47811: gap of unknown length
            *    47812    49766: contig of 1955 bp in length
            *    49767    49866: gap of unknown length
            *    49867    53471: contig of 3605 bp in length
            *    53472    53571: gap of unknown length
            *    53572    55697: contig of 2126 bp in length
            *    55698    55797: gap of unknown length
            *    55798    58869: contig of 3072 bp in length
            *    58870    58969: gap of unknown length
            *    58970    62248: contig of 3279 bp in length
            *    62249    62348: gap of unknown length
            *    62349    65436: contig of 3088 bp in length
            *    65437    65536: gap of unknown length
            *    65537    67322: contig of 1786 bp in length
            *    67323    67422: gap of unknown length
            *    67423    69569: contig of 2147 bp in length
            *    69570    69669: gap of unknown length
```

```
            *    69670    72310: contig of 2641 bp in length
            *    72311    72410: gap of unknown length
            *    72411    75766: contig of 3356 bp in length
            *    75767    75866: gap of unknown length
            *    75867    77612: contig of 1746 bp in length
            *    77613    77712: gap of unknown length
            *    77713    80441: contig of 2729 bp in length
            *    80442    80541: gap of unknown length
            *    80542    82037: contig of 1496 bp in length
            *    82038    82137: gap of unknown length
            *    82138    84137: contig of 2000 bp in length
            *    84138    84237: gap of unknown length
            *    84238    85983: contig of 1746 bp in length
            *    85984    86083: gap of unknown length
            *    86084    88084: contig of 2001 bp in length
            *    88085    88184: gap of unknown length
            *    88185    89469: contig of 1285 bp in length
            *    89470    89569: gap of unknown length
            *    89570    90925: contig of 1356 bp in length
            *    90926    91025: gap of unknown length
            *    91026    92114: contig of 1089 bp in length
            *    92115    92214: gap of unknown length
            *    92215    93366: contig of 1152 bp in length
            *    93367    93466: gap of unknown length
            *    93467    94962: contig of 1496 bp in length
            *    94963    95062: gap of unknown length
            *    95063    96516: contig of 1454 bp in length
            *    96517    96616: gap of unknown length
            *    96617    98564: contig of 1948 bp in length
            *    98565    98664: gap of unknown length
            *    98665   100193: contig of 1529 bp in length
            *   100194   100293: gap of unknown length
            *   100294   101520: contig of 1227 bp in length
            *   101521   101620: gap of unknown length
            *   101621   102748: contig of 1128 bp in length
            *   102749   102848: gap of unknown length
            *   102849   104031: contig of 1183 bp in length
            *   104032   104131: gap of unknown length
            *   104132   105772: contig of 1641 bp in length.
FEATURES             Location/Qualifiers
     source          1. .105772
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-10E14"
BASE COUNT    24796 a  25921 c  26128 g  24930 t   3997 others
ORIGIN


  Query Match             8.3%;  Score 156.6;  DB 2;  Length 105772;
  Best Local Similarity   85.7%;  Pred. No. 3.4e-17;
  Matches  174;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy     847 cgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggacagagca 906
           | |||| ||   ||||||||||||  | || ||||||||||||| |||||||||| ||
Db   21223 CCCAGACCCTGGGGACTGGAGCCCTGGTAACGTGCAGAAGTGGCTTCTGTGGACAGAACA 21282

Qy     907 ccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaaggagctgtg 966
           ||| |||||||||| ||   ||||||||||||||||||||||  ||||| ||||| ||
Db   21283 CCAGTACCGGCTGCCTCCAGCGGGCAAGGCCTTCCAGGAGCTGGGTGGCAAAGAGCTATG 21342

Qy     967 cgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgctgcacgc 1026
           ||||||||| |||||||||||||||||||  | ||| |||| ||||| |||||||| ||
Db   21343 CGCCATGTCCGAGGAGCAGTTCCGCCAGCGGGCACCCTTGGGCGGGGACGTGCTGCATGC 21402

Qy    1027 ccacctggacatctggaagtcag 1049
           |||||||||||||||||||||||
Db   21403 TCACCTGGACATCTGGAAGTCAG 21425


Search completed: August  6, 2002, 19:53:16
Job time: 5825 sec
```

GenCore version 4.5 .

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:52:21 ; Search time 311.57 Seconds
(without alignments)
10436.935 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1870	98.7	1894	21	AAZ50691	Human prostate der
	2	1870	98.7	1894	22	AAH47005	Human ETS suppress
	3	1870	98.7	1894	22	AAC83261	Gland-specific Ets
	4	1868.4	98.6	1918	22	AAD07337	Human transcriptio
	5	1865.2	98.5	3317	22	AAC83266	Gland-specific Ets
	6	1861.4	98.3	1905	22	AAH42468	Nucleotide sequenc
	7	1018.6	53.8	1087	21	AAF21828	Human breast and o
	8	738.4	39.0	1704	22	AAH47006	Mouse ETS suppress
	9	391	20.6	392	21	AAH30613	Human colon cancer
	10	350	18.5	427	21	AAC77746	Human cancer assoc
	11	329.8	17.4	1751	22	AAC83263	GSEF promoter and
	12	263	13.9	278	21	AAA06617	Human immunogenic
	13	263	13.9	278	22	AAS63826	Human cDNA encodin
	14	263	13.9	278	22	AAH93733	Human prostate-spe
	15	263	13.9	278	22	AAH85047	Human prostate-spe
	16	263	13.9	278	22	AAH02798	Prostate tumour an
	17	237	12.5	264	21	AAC04502	Human secreted pro
	18	147.6	7.8	1515	23	ABL09673	Drosophila melanog
	19	147.6	7.8	5325	23	ABL09672	Drosophila melanog
	20	96	5.1	5317	23	ABL10827	Drosophila melanog
	21	96	5.1	18727	23	ABL10826	Drosophila melanog
	22	83.2	4.4	2266	20	AAZ41049	Human ELK-1 encodi
	23	83.2	4.4	2266	20	AAZ06608	Human ELK-1 nucleo
	24	80.4	4.2	2286	24	ABI99696	Mouse ischaemic co
	25	75.2	4.0	2667	18	AAT47198	Human ETS2 repress
	26	75	4.0	1116	19	AAV38354	Transcription fact
	27	75	4.0	1116	21	AAZ49125	Human ESX coding s
	28	75	4.0	1907	19	AAV38355	Transcription fact
	29	75	4.0	1907	21	AAZ36707	Nucleotide sequenc
	30	75	4.0	1995	22	AAF98700	Human ovarian canc
	31	75	4.0	1996	21	AAC77813	Human cancer assoc
	32	75	4.0	2266	20	AAZ33671	Human breast tumou
c	33	74.8	3.9	567	22	ABA48124	Human breast cell
c	34	74.8	3.9	567	22	ABA66003	Human foetal liver
c	35	74.8	3.9	567	22	ABA33090	Probe #11556 for g
c	36	74.8	3.9	567	22	AAK14426	Human brain expres
c	37	74.8	3.9	567	22	AAK40160	Human bone marrow
c	38	74.8	3.9	567	22	AAI20932	Probe #10865 for g
c	39	74.8	3.9	567	22	AAI46176	Probe #14862 used
c	40	74.8	3.9	567	22	AAI06643	Probe #6634 used t
	41	74.8	3.9	2064	17	AAT37087	E1AF matrix metall
	42	74.8	3.9	2064	21	AAZ50016	Human polyomavirus
	43	74.8	3.9	2073	15	AAQ55149	Adenovirus E1A-F g
c	44	74	3.9	473	22	ABA50193	Human breast cell
c	45	74	3.9	473	22	ABA68128	Human foetal liver

ALIGNMENTS

RESULT 1
AAZ50691
ID AAZ50691 standard; cDNA; 1894 BP.
XX
AC AAZ50691;
XX
DT 23-MAY-2000 (first entry)
XX
DE Human prostate derived Ets factor cDNA.
XX
KW Prostate Derived Ets factor; PDEF; human; chromosome 6p21.3; cancer;
KW loss of heterozygosity; chromosomal translocation; linkage analysis;
KW cytostatic; cardiant; immunosuppressive; cerebroprotective; fungicide;
KW antibacterial; vulnerary; neuroprotective; antiparkinsonian; nootropic;
KW anabolic; antiinflammatory; anorectic; hybridisation probe; forensic;
KW tumour marker; diagnosis; treatment; prostate cancer; blood coagulation;
KW autoimmune disorder; haematopoietic; immune/nervous system; stroke;
KW neoplasm; microbial infection; tissue regeneration; heart attack;
KW scarring; food additive; preservative; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 416..1423
FT /*tag= a
FT /product= "Human prostate derived Ets factor"
FT /note= "Expressed in prostate epithelium"
XX
PN WO200006589-A1.
XX
PD 10-FEB-2000.
XX

PF 02-AUG-1999; 99WO-US17470.
XX
PR 31-JUL-1998; 98US-0126945.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Libermann TA, Oettgen JP, Kunsch CA, Endress GA, Rosen CA;
XX
DR WPI; 2000-195255/17.
DR P-PSDB; AAY45003.
XX
PT Novel prostate derived polypeptide, polynucleotide useful for
PT diagnosis, prevention and treatment of prostate cancer, autoimmune
PT disorders, microbial infections and also as food additive or
PT preservative -
XX
PS Claim 4; Fig 1; 132pp; English.
XX
CC The present cDNA sequence encodes a human prostate derived Ets factor
CC (PDEF), isolated from human prostate carcinoma cDNA library
CC (ATCC No.203072). PDEF gene expression is highest in tissues involved in
CC androgen sensitivity, with enriched expression in prostate epithelium
CC cells. PDEF gene is mapped to the human chromosome 6p21.3
CC region that is associated with loss of heterozygosity and chromosomal
CC translocations in various human cancers. PDEF has cytostatic, cardiant,
CC immunosuppressive, cerebroprotective, fungicide, antibacterial,
CC vulnerary, neuroprotective, antiparkinsonian, nootropic, anabolic,
CC antiinflammatory and anorectic activity. PDEF polynucleotides are useful
CC in linkage analysis as markers, as hybridisation probes for differential
CC identification of the tissues or cell types and as polymorphic markers
CC for forensic purposes. PDEF is useful as prostate-specific tumour marker
CC for the diagnosis and treatment of prostate cancer. PDEF sequences are
CC useful for treating autoimmune disorders, haematopoietic, blood
CC coagulation, immune and nervous system disorders, hyperproliferative
CC disorders like, neoplasms and microbial infections, heart attacks,
CC stroke, scarring and for tissue regeneration. They are also useful as
CC food additives or preservatives.
XX
SQ Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;

Query Match 98.7%; Score 1870; DB 21; Length 1894;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccacccctttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
             |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tcagagggccacccctttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db     661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
```

```
Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

RESULT 2
AAH47005
ID AAH47005 standard; cDNA; 1894 BP.
XX
AC AAH47005;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human ETS suppressor factor (esf) protein encoding cDNA.
XX
KW ETS suppressor factor; esf; apoptosis; cancer; cytostatic; human;
KW growth-suppressor; transcription factor; gene-therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 416..1423
FT /*tag= a
FT /product= "human esf"
XX
PN WO200157186-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03640.
XX
PR 04-FEB-2000; 2000US-0180244.
XX
PA (MUSC-) MUSC FOUND RES DEV.
XX
PI Watson DK;
XX
DR WPI; 2001-465703/50.
DR P-PSDB; AAB85609.
XX
PT Fragment of a human ETS suppressor factor protein for inhibiting the
PT growth or proliferation of and/or inducing apoptosis of cancer cells,
PT especially prostate, breast, colon or cervical cancer cells for
PT treatment -
XX
PS Examples; Fig 1A; 94pp; English.
XX
CC The invention relates to methods for treating or preventing cancer by
CC modifying the expression of ETS suppressor factor (esf) protein gene or
CC the activity of the gene product. The esf protein or its modified form
CC are useful for inhibiting the growth or proliferation of and/or inducing
CC apoptosis of cancer cells, especially prostate, breast, colon or cervical
CC cancer cells. They are useful for treating cancer in a subject when
CC introduced in to the vicinity of the cancer in a subject. The methods
CC provided are useful for inhibiting the specific degradation of esf
CC protein in a cell; for diagnosing a preneoplastic or neoplastic condition
CC in a human by e.g. PCR and for identifying a compound that inhibits the
CC degradation of esf protein in cancer cells. The present sequence
CC represents a human esf protein encoding cDNA.
XX
SQ Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;

```
  Query Match            98.7%;  Score 1870;  DB 22;  Length 1894;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy    181 gccaccagccacccctgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
            |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tcagagggccacccctgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgacac 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
```

```
Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

```
RESULT   3
AAC83261
ID   AAC83261 standard; DNA; 1894 BP.
XX
AC   AAC83261;
XX
DT   16-MAR-2001  (first entry)
XX
DE   Gland-specific Ets transcription factor (GSEF) nucleotide sequence.
XX
KW   Transcription factor; gland-specific Ets transcription factor; GSEF;
KW   metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW   leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3; ds.
XX
OS   Homo sapiens.
XX
PN   WO200070092-A1.
XX
PD   23-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US13173.
XX
PR   14-MAY-1999;   99US-0134112.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kaufmann J,  Xin H,  Harrowe G;
XX
DR   WPI; 2001-041019/05.
DR   P-PSDB; AAB49628.
XX
PT   Detecting metastatic and potential metastatic cancerous cells, useful
PT   for diagnosing, prognosing, grading and staging of cancers by detecting
PT   gland-specific Ets transcription factor gene product in a biological
PT   sample from a cell  -
XX
PS   Disclosure; Page 81-83; 95pp; English.
XX
CC   This invention relates to a method for the detection and determination of
CC   the metastatic potential of a cell. The method comprises detecting a
CC   gland-specific Ets transcription factor (GSEF) gene product in a test
CC   sample. Detection of a GSEF gene product in the test sample in amount
CC   lower than that in a normal cell, is indicative of a cell with high
CC   metastatic potential. The method is useful for determining the metastatic
CC   potential of a cell, for the diagnosis and prognosis of cancer as well as
CC   grading and staging of cancers by detecting GSEF expression in a
CC   biological test sample. The method may also be used to monitor patients
CC   having a predisposition to develop a particular cancer. GSEF polypeptides
CC   are useful for producing antibodies, in cancer diagnosis, prognosis,
CC   grading, staging and management of breast and prostate tumours, and in
CC   detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC   useful in gene therapy. GSEF gene product expression levels can be used
CC   in conjunction with any tissue in which an alteration in GSEF gene
CC   product expression levels is associated with development of a
CC   cancer-associated phenotype. Cancers, which can be monitored include
CC   cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC   adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC   lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC   human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC   represents the DNA encoding GSEF.
XX
SQ   Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;
```

```
  Query Match             98.7%;  Score 1870;  DB 22;  Length 1894;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
```

```
Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
           |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tcagagggccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db     661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy    1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800

Qy    1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy    1861 cccgggaatggataataaagatactagagaactg 1894
           |||||||||||||||||||||||||||||||||||
Db    1861 cccgggaatggataataaagatactagagaactg 1894
```

RESULT 4
AAD07337
ID AAD07337 standard; cDNA; 1918 BP.
XX
AC AAD07337;
XX
DT 04-AUG-2001 (first entry)
XX
DE Human transcription factor, prost-ets cDNA.
XX
KW Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;
KW vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;
KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 422..1429
FT /*tag= a
FT /product= "Human PROST-Ets protein"
XX

PN WO200142472-A2.
XX
PD 14-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US42295.
XX
PR 30-NOV-1999; 99US-0168182.
PR 21-NOV-2000; 2000US-0718159.
XX
PA (SCHD) SCHERING AG.
XX
PI Harkins R, Lin RJ, Luke M, Monteclaro F, Parkes D, Parry G;
PI Steinbrecher R, Van Heuit PT, Xuan J;
XX
DR WPI; 2001-381692/40.
DR P-PSDB; AAE03556.
XX
PT New transcription factor polypeptides, designated PROST-Ets
PT polypeptides, useful for diagnostic, research and therapeutic
PT applications -
XX
PS Claim 8; Fig 1; 74pp; English.
XX
CC The invention relates to human transcription factor protein, designated
CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and
CC nucleic acids are useful for treating a disease-state in human patient,
CC such as prostate, breast and ovarian cancer which is associated with
CC inappropriate expression of PROST-Ets and where the patient is in need of
CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for
CC diagnosing and treating diseases of cell proliferation, such as prostate
CC cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used
CC as DNA probes, as target for antisense and ribozyme therapy, as template
CC for producing antisense polynucleotides, and as diagnostic reagent for
CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is
CC also used in gene therapy. PROST-Ets protein is useful for generating
CC antibodies, and for stimulating immune response in cells containing
CC PROST-Ets proteins. The present cDNA sequence encodes human transcription
CC factor, designated as PROST-Ets. Prost-ets cDNA was found in a cDNA clone
CC obtained from human prostate tissues.
XX
SQ Sequence 1918 BP; 385 A; 656 C; 574 G; 303 T; 0 other;

Query Match 98.6%; Score 1868.4; DB 22; Length 1918;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 66

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 126

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 186

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
           |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 246

Qy    241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 306

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    307 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 366

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    367 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 426

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 486

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 546

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    547 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 606

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    607 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 666

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    667 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 726

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    727 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 786

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    787 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 846

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    847 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 906

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    907 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 966

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    967 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1026

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1086

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1146

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1206

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1266

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1267 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1326

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1327 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1386

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1387 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1446

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1447 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1506
```

```
Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1507 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1566

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   1567 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcaggg 1626

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1627 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1686

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1687 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1746

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1747 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1806

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1807 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1866

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1867 cccgggaatggataataaagatactagagaactg 1900
```

RESULT 5
AAC83266
ID AAC83266 standard; DNA; 3317 BP.
XX
AC AAC83266;
XX
DT 16-MAR-2001 (first entry)
XX
DE Gland-specific Ets transcription factor (GSEF) cDNA sequence.
XX
KW Transcription factor; gland-specific Ets transcription factor; GSEF;
KW metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3; ss.
XX
OS Homo sapiens.
XX
PN WO200070092-A1.
XX
PD 23-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US13173.
XX
PR 14-MAY-1999; 99US-0134112.
XX
PA (CHIR) CHIRON CORP.
XX
PI Kaufmann J, Xin H, Harrowe G;
XX
DR WPI; 2001-041019/05.
DR P-PSDB; AAB49633.
XX
PT Detecting metastatic and potential metastatic cancerous cells, useful
PT for diagnosing, prognosing, grading and staging of cancers by detecting
PT gland-specific Ets transcription factor gene product in a biological
PT sample from a cell -
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC This invention relates to a method for the detection and determination of
CC the metastatic potential of a cell. The method comprises detecting a
CC gland-specific Ets transcription factor (GSEF) gene product in a test
CC sample. Detection of a GSEF gene product in the test sample in amount
CC lower than that in a normal cell, is indicative of a cell with high
CC metastatic potential. The method is useful for determining the metastatic
CC potential of a cell, for the diagnosis and prognosis of cancer as well as
CC grading and staging of cancers by detecting GSEF expression in a
CC biological test sample. The method may also be used to monitor patients
CC having a predisposition to develop a particular cancer. GSEF polypeptides
CC are useful for producing antibodies, in cancer diagnosis, prognosis,
CC grading, staging and management of breast and prostate tumours, and in
CC detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC useful in gene therapy. GSEF gene product expression levels can be used
CC in conjunction with any tissue in which an alteration in GSEF gene
CC product expression levels is associated with development of a
CC cancer-associated phenotype. Cancers, which can be monitored include
CC cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC represents the cDNA encoding GSEF.
XX
SQ Sequence 3317 BP; 710 A; 1026 C; 970 G; 611 T; 0 other;

Query Match 98.5%; Score 1865.2; DB 22; Length 3317;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1400 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 1459

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1460 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 1519

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1520 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 1579

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1580 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 1639

Qy    241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1640 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 1699

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1700 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 1759

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1760 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 1819

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1820 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 1879

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1880 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 1939

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1940 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 1999

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2000 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 2059

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db   2060 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 2119
```

```
Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2120 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 2179

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2180 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 2239

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2240 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 2299

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2300 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 2359

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2360 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggagatgtgct 2419

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2420 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 2479

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2480 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 2539

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 2599

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2600 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 2659

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2660 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 2719

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   2720 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 2779

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 2839

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 2899

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2900 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 2959

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2960 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcaggg 3019

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3020 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 3079

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3080 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 3139

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3140 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 3199

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   3200 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 3259

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||| ||||||||||||||||
Db   3260 cccgggaatggataatagagatactagagaactg 3293


RESULT   6
AAH42468
ID   AAH42468 standard; cDNA; 1905 BP.
XX
AC   AAH42468;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of human prostate associated Ets protein (PRAEP).
XX
KW   Human; prostate associated Ets protein; PRAEP; immunological disorder;
KW   cell proliferation disorder; reproductive disorder; psoriasis; diabetes;
KW   developmental disorder; cancer; arteriosclerosis; multiple sclerosis;
KW   acquired immunodeficiency syndrome; AIDS; polycystic ovary syndrome;
KW   ectopic pregnancy; gonadal dysgenesis; achondroplastic dwarfism; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             430..1437
FT                   /*tag= a
FT                   /product= "prostate associated Ets protein (PRAEP)"
XX
PN   US6265565-B1.
XX
PD   24-JUL-2001.
XX
PF   03-APR-1998;   98US-0055113.
XX
PR   03-APR-1998;   98US-0055113.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Bandman O,  Guegler KJ,  Lal P,  Corley NC;
XX
DR   WPI; 2001-463945/50.
DR   P-PSDB; AAG63017.
XX
PT   Polynucleotides encoding human prostate associated Ets proteins, useful
PT   for preventing, diagnosing and treating, e.g. cancer, diabetes,
PT   psoriasis and dwarfism -
XX
PS   Claim 3; Fig 1A-E; 31pp; English.
XX
CC   The present sequence encodes a human prostate associated Ets protein
CC   (PRAEP). PRAEP may be used in the prevention, treatment and diagnosis
CC   of diseases associated with inappropriate PRAEP expression such as
CC   cell proliferation, immunological, reproductive and developmental
CC   disorders. The PRAEP polypeptides may be used as antigens in the
CC   production of antibodies against PRAEP and in assays to identify
CC   modulators (agonists and antagonists) of PRAEP expression and activity.
CC   The anti-PRAEP antibodies and PRAEP antagonists may also be used to
CC   down regulate PRAEP expression and activity. The anti-PRAEP antibodies
CC   and antagonists may be used to prevent and treat cell proliferation
CC   disorders (e.g. cancer, arteriosclerosis and psoriasis), immunological
CC   disorders (e.g. acquired immunodeficiency syndrome (AIDS), multiple
CC   sclerosis and diabetes), reproductive disorders (e.g. polycystic ovary
CC   syndrome and ectopic pregnancy) and disorders of growth and development
CC   (e.g. gonadal dysgenesis and achondroplastic dwarfism).
XX
SQ   Sequence 1905 BP; 365 A; 661 C; 576 G; 303 T; 0 other;


  Query Match             98.3%;  Score 1861.4;  DB 22;  Length 1905;
  Best Local Similarity   99.2%;  Pred. No. 0;
```

```
  Matches 1871;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 74

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 134

Qy    121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 194

Qy    181 gccaccagccacccctgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 tcagagggccacccctgagggtggccaggcccccagtggccaacctgagtgctgcctct 254

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 314

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    315 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 374

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 434

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 494

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 554

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    555 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 614

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 674

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    675 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 734

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 794

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 854

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 914

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 974

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1034

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1094

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1095 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1154

Qy.  1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1155 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1214

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1215 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1274

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1275 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1334

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1335 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1394

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1395 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1454

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1455 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1514

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1515 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1574

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1575 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcaggg 1634

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1635 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1694

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1695 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1754

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1755 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1814

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1815 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1874

Qy   1861 cccgggaatggataataaagatactag 1887
          |||||||||||||||||||||||||||
Db   1875 cccgggaatggataataaagatactag 1901


RESULT   7
AAF21828
ID   AAF21828 standard; DNA; 1087 BP.
XX
AC   AAF21828;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Human breast and ovarian cancer associated antigen gene SEQ ID 215.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
```

```
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   P-PSDB; AAB58925.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 1; Page 642-643; 1299pp; English.
XX
CC   Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC   proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC   associated with breast and ovarian cancer. Included in the invention are
CC   sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC   isolation and characterisation of the DNA and protein sequences of the
CC   invention. The breast and ovarian cancer associated DNA, protein, agonist
CC   or antagonist sequences exhibit cytostatic; immunosuppressive;
CC   nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC   antibacterial; antifungal; antiparasitic and cardiant activity. The
CC   polynucleotide and protein sequences are used in the diagnosis of cancer,
CC   particularly breast and ovarian cancer. The nucleic acid sequences,
CC   proteins, agonists and agonists may also be used in the diagnosis,
CC   prevention and treatment of immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; cardiovascular disorders such as
CC   myocardial ischaemias; wound healing; neurological diseases such as
CC   cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ   Sequence 1087 BP; 215 A; 347 C; 339 G; 182 T; 4 other;

  Query Match             53.8%;  Score 1018.6;  DB 21;  Length 1087;
  Best Local Similarity   97.0%;  Pred. No. 4.8e-205;
  Matches 1055;  Conservative    3;  Mismatches   28;  Indels    2;  Gaps    2;

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      2 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 61

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 121

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    122 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 181

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 241

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    242 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 301

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 361

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 421

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 481

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    482 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 541

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    542 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 601

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    602 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 661

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    662 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 721

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          |||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||
Db    722 ggactcagcccaggtggcccggctgtrgggcatccgcaagaaccgtcccgccatgaacta 781

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||:||||  ||||||||||||||||||||||||||||||
Db    782 cgacaagctgagccgctccatccgscagtattacaagaagggcatcatccggaagccaga 841

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          |||||  ||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db    842 catctgccagcgsctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 901

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    902 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 961

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||| |||||| |||||||||||||||||| |||||||||||||| ||||||
Db    962 ggggaaaac-ggcagtctgctctgctgctctgaccct-cagagcccaaggtcaaggaggg 1019

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||  ||||     ||  |   ||||||||||||||| |||||    |
Db   1020 gcaaccaactgcccaggggatatgggtcctcttggggccttcgggaccctggggncaag 1079

Qy   1621 gtgcttcc 1628
          | |||| |
Db   1080 gggctttc 1087


RESULT   8
AAH47006
ID   AAH47006 standard; cDNA; 1704 BP.
XX
AC   AAH47006;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Mouse ETS suppressor factor (esf) protein encoding cDNA.
XX
KW   ETS suppressor factor; esf; apoptosis; cancer; cytostatic; mouse;
KW   growth-suppressor; transcription factor; gene-therapy; ss.
XX
OS   Mus sp.
```

```
XX
FH   Key             Location/Qualifiers
FT   CDS             362..1339
FT                   /*tag=  a
FT                   /product= "mouse esf"
XX
PN   WO200157186-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03640.
XX
PR   04-FEB-2000; 2000US-0180244.
XX
PA   (MUSC-) MUSC FOUND RES DEV.
XX
PI   Watson DK;
XX
DR   WPI; 2001-465703/50.
DR   P-PSDB; AAB85610.
XX
PT   Fragment of a human ETS suppressor factor protein for inhibiting the
PT   growth or proliferation of and/or inducing apoptosis of cancer cells,
PT   especially prostate, breast, colon or cervical cancer cells for
PT   treatment -
XX
PS   Examples; Fig 1C; 94pp; English.
XX
CC   The invention relates to methods for treating or preventing cancer by
CC   modifying the expression of ETS suppressor factor (esf) protein gene or
CC   the activity of the gene product. The esf protein or its modified form
CC   are useful for inhibiting the growth or proliferation of and/or inducing
CC   apoptosis of cancer cells, especially prostate, breast, colon or cervical
CC   cancer cells. They are useful for treating cancer in a subject when
CC   introduced in to the vicinity of the cancer in a subject. The methods
CC   provided are useful for inhibiting the specific degradation of esf
CC   protein in a cell; for diagnosing a preneoplastic or neoplastic condition
CC   in a human by e.g. PCR and for identifying a compound that inhibits the
CC   degradation of esf protein in cancer cells. The present sequence
CC   represents a mouse esf protein encoding cDNA.
XX
SQ   Sequence 1704 BP; 369 A; 527 C; 491 G; 317 T; 0 other;


  Query Match             39.0%;  Score 738.4;  DB 22;  Length 1704;
  Best Local Similarity   73.3%;  Pred. No. 4.4e-146;
  Matches 1061;  Conservative    0;  Mismatches  336;  Indels   51;  Gaps    7;

Qy      38 tgtccacactgccccacagacccagtcctccaagcctgctgccagctccctgcaagcccc 97
           ||||| | ||||| | || ||| ||  | |||  ||      | || ||    |   |||
Db       1 tgtccgctctgcctcccacacctagcaccccagccc--gctgctgccccggtgagaaccc 58

Qy      98 tcaggttgggccttgccacggtgccagcaggcagccctgggctgggggtaggggactccc 157
              ||| |||||||||| || ||||||||||||  |||||| |||   |  ||||| || ||
Db      59 ccagcttgggccttgtcatggtgccagcaggtggccctgagcttctgacaggggcctgcc 118

Qy     158 tacaggcacgcagccctgagaccgccaccagccacccctgagggtggccaggccccccag 217
           || || |  |||| |||||| ||     ||| | | |  | | || |     | |||  |
Db     119 tatagacctgcaggcctgaggcc----tcagactcacactcaaggggcaagaggccctgg 174

Qy     218 tggccaacctgagtgctgcctctgccaccagccctgctggcccctggttccgctggcccc 277
           ||||| |||| || ||  |||||| | ||||||||||||| |||   | |      | |
Db     175 tggcccacctaagagccacctctgtccccagccctgctgccccactgat------gtctg 228

Qy     278 ccagatgcctggctgagacacgccagtggcctcagctgcccacacctcttcccggccccct 337
           | ||  ||  || | ||| |     ||  ||      ||    |||| ||| || ||||
Db     229 actgagacccagcagtgaccc-----tgagctgcctgcccactgcctcctcctggtccct 283

Qy     338 ggagttggcactgcagcagacagctccctgggcaccaggcagctaacagacacagccgcc 397
           |  |||||| |||| |  ||| |    ||   |   ||   ||    |     ||| |||
Db     284 gaggttggctctgccgaggacggacgactcttctgaagcaggcggctaacggaagcagcc 343

Qy     398 agcccaaacagcagcggcatgggcagcgccagcccgggtctgagcagcgtatcccccagc 457
```

```
Db     344 ccaagcctccaccgcagcatgggcagtgccagcccaggcctgagcaacgtgtcccccggt 403

Qy     458 cacctcctgctgccccccgacacggtgtcgcggacaggcttggagaaggcggcagcgggg 517
             ||| || |||  ||| ||   ||   | || |||||   |||||||||| ||| | ||
Db     404 tgcctgctactgttcccagatgtggcaccacgaacagggacggagaaggcagcatcagga 463

Qy     518 gcagtgggtctcgagagacgggactggagtcccagtccacccgccacgcccgagcagggc 577
           ||| |||| |  ||||  | ||| |||||||| |||||||||||||| || |||||||||
Db     464 gcaatgggccctgagaagcaggaatggagtcctagtccacccgccacccctgagcagggc 523

Qy     578 ctgtccgccttcgacctctcctactttgacatgctgtaccctgaggacagcagctgggca 637
           ||||| || ||| ||||||| |||||| |||   |||| || || || ||||||||||
Db     524 ctgtctgctttctacctctcttactttaaca---tgtatcccgacgatagcagctgggtc 580

Qy     638 gccaaggcccctggggccagcagtcgggaggagccacctgaccagcctgagcagtgcccg 697
           ||||| | ||| | |||| |     |||||||  |  || ||  |||| |||||||| ||
Db     581 gccaaagtccccgaggcccgtgccggggaggaccacccggaggagcccgagcagtgtccc 640

Qy     698 gtcattgacagccaagccccagcgggcaccctggacttggtgcccggcgggctgaccttg 757
           |||||||||||||| ||| | | | ||                            || |||
Db     641 gtcattgacagccaggcctctgggagc----------------------------acgttg 673

Qy     758 gaggagcactcgctggagcaggtgcagtccatggtggtgggcgaagtgctcaaggacatc 817
           || |||||||||| |||||||||| || ||||| |||||||| || || || || ||
Db     674 gatgagcactcgctagagcaggtgcaatcgatggttgtgggcgaggtcctgaaagatatt 733

Qy     818 gagacggcctgcaagctgctcaacatcaccgcagatcccatggactggagccccagcaat 877
           ||||||||||||||| || |||||||| ||||| ||   ||||||||||||  | ||
Db     734 gagacggcctgcaagcttctgaacatcacagcagaccctggggactggagccctggtaac 793

Qy     878 gtgcagaagtggctcctgtggacagagcaccaataccggctgccccccatgggcaaggcc 937
           ||||||||||||||  | |||||||| ||||| |||||||||| ||    |||||||||
Db     794 gtgcagaagtggcttttatggacagaacaccagtaccggctgcctccagcaggcaaggcc 853

Qy     938 ttccaggagctggcgggcaaggagctgtgcgccatgtcggaggagcagttccgccagcgc 997
           |||||||||||||  || ||||||||||||||||||| ||||| |||||||| |||||
Db     854 ttccaggagctgggcggtaaggagctgtgcgccatgtccgaggaacagttccgtcagcgt 913

Qy     998 tcgcccctgggtggggatgtgctgcacgcccacctggacatctggaagtcagcggcctgg 1057
            | ||| |||||||||||||| ||||| ||||||||||||||||||||||||||||||||
Db     914 gcacccttgggtggggatgtactgcatgcccacctggacatctggaagtcagcggcctgg 973

Qy    1058 atgaaagagcggacttcacctggggcgattcactactgtgcctcgaccagtgaggagagc 1117
           ||||| ||| |||| || |||||| |  |||||||||| ||||| ||||| |||||| ||
Db     974 atgaaggagaggacctcgcctgggacccttcactactgcgcctccaccagcgaggagggc 1033

Qy    1118 tggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctgtggcagttc 1177
           ||||| ||  | ||||||||||| || ||||||||||||||||| ||||||||||||||
Db    1034 tggacggatggtgaggtggactcgtcgtgctccgggcagcccattcacctgtggcagttc 1093

Qy    1178 ctcaaggagttgctactcaagccccacagctatggccgcttcattaggtggctcaacaag 1237
           || || ||  |||| ||||||||||||||||||||||||||||  | ||||||||||||
Db    1094 ctgaaagaactgctgctcaagccccacagctatggccgcttcatccgctggctcaacaag 1153

Qy    1238 gagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgc 1297
           ||||| |||||||||||||||||||||||||||| |||||||||| ||||||||  | |||
Db    1154 gagaaaggcatcttcaaaattgaggactcagcacaggtggcccgactgtggggtgtgcgc 1213

Qy    1298 aagaaccgtcccgccatgaactgcgacaagctgagccgctccatccgccagtcttacaag 1357
           |||||||| || ||||||||||  || || || ||||||||||||||||||| |||||||
Db    1214 aagaaccggccagccatgaactatgataaactaagccgctccatccgccagtattacaag 1273

Qy    1358 aagggcatcatccggaagccagacatctcccagcgcctcgtctaccagttcgtgcacccc 1417
           |||||||||||| || || || |||||||| |||||||| |||||||| || ||||| ||
Db    1274 aagggcatcattcgtaaacccgacatctctcagcgccttgtctaccaatttgtgcatcca 1333

Qy    1418 atctgagtgcc----tggcccagggcctgaaacccgccctcaggggcctctctcctgcct 1473
           ||||||  |||     | ||   |||||  |||| ||||   |  ||| |||||  |
Db    1334 gtctgagagccacagagaccagaggcctacaacctgccccaggcagccactctctggttg 1393

Qy    1474 gccctgcc 1481
           |||  | |
```

```
Db   1394 gcctggtc 1401

RESULT  9
AAH30613
ID   AAH30613 standard; cDNA; 392 BP.
XX
AC   AAH30613;
XX
DT   27-JUL-2001  (first entry)
XX
DE   Human colon cancer cell line Km12L4-A cDNA library derived sequence #547.
XX
KW   Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
KW   detection; colon cancer cell line Km12L4-A; ss.
XX
OS   Homo sapiens.
XX
PN   WO200018916-A2.
XX
PD   06-APR-2000.
XX
PF   23-SEP-1999;   99WO-US22226.
XX
PR   28-SEP-1998;   98US-0102161.
PR   28-SEP-1998;   98US-0102180.
PR   29-SEP-1998;   98US-0102380.
PR   08-OCT-1998;   98US-0103815.
PR   27-OCT-1998;   98US-0105877.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Sudduth-Klinger J;
PI   Reinhard C,  Giese K,  Randazzo F,  Kennedy GC,  Pot D,  Kassam A;
PI   Lamson G,  Drmanac R,  Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I;
PI   Leshkowitz D,  Kita D,  Garcia V,  Jones LW,  Stache-Crain B;
XX
DR   WPI; 2000-293155/25.
XX
PT   Polynucleotide library comprising 1079 defined sequences,  useful in
PT   the form of an array to detect cancer or susceptibility to cancer -
XX
PS   Claim 1; Page 340; 502pp; English.
XX
CC   The present invention describes a library of polynucleotides comprising
CC   1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
CC   are: (1) an isolated polynucleotide (I) having at least 90% identity to
CC   one of the 1079 sequences; (2) a recombinant host cell containing (I);
CC   (3) an isolated polypeptide (II) encoded by (I); (4) an antibody that
CC   specifically binds to (II); (5) a vector comprising (I); and (6) a method
CC   of detecting differentially expressed genes correlated with a cancerous
CC   state of a mammalian cell comprising detecting a gene product encoded by
CC   65 of the 1079 sequences given in the specification. The polynucleotides
CC   are used to monitor patients having (or susceptible) to cancer to detect
CC   potentially malignant events at a molecular level before they are
CC   detectable at a gross morphological level. The polynucleotides are also
CC   useful for monitoring the efficacy of various therapies and preventive
CC   interventions. Polynucleotide probes based on the disclosed sequences
CC   are useful for chromosome mapping and detection of transcription levels.
CC   The 1079 polynucleotide sequences were derived from a human colon cancer
CC   cell line Km12L4-A cDNA library.
XX
SQ   Sequence 392 BP; 82 A; 119 C; 126 G; 65 T; 0 other;


  Query Match             20.6%;  Score 391;  DB 21;  Length 392;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-73;
  Matches  391;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    799 cgaagtgctcaaggacatcgagacggcctgcaagctgctcaacatcaccgcagatcccat 858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 cgaagtgctcaaggacatcgagacggcctgcaagctgctcaacatcaccgcagatcccat 60

Qy    859 ggactggagccccagcaatgtgcagaagtggctcctgtggacagagcaccaataccggct 918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ggactggagccccagcaatgtgcagaagtggctcctgtggacagagcaccaataccggct 120

Qy    919 gccccccatgggcaaggccttccaggagctggcgggcaaggagctgtgcgccatgtcgga 978
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gccccccatgggcaaggccttccaggagctggcgggcaaggagctgtgcgccatgtcgga 180

Qy    979 ggagcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacat 1038
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ggagcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacat 240

Qy   1039 ctggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgc 1098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ctggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgc 300

Qy   1099 ctcgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcc 1158
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ctcgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcc 360

Qy   1159 catccacctgtggcagttcctcaaggagttg 1189
          |||||||||||||||||||||||||||||||
Db    361 catccacctgtggcagttcctcaaggagttg 391


RESULT  10
AAC77746
ID   AAC77746 standard; cDNA; 427 BP.
XX
AC   AAC77746;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:140.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antianglogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43537.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 720-721; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
```

```
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 427 BP; 73 A; 161 C; 119 G; 63 T; 11 other;


  Query Match             18.5%;  Score 350;  DB 21;  Length 427;
  Best Local Similarity   92.0%;  Pred. No. 1.6e-64;
  Matches  390;  Conservative    9;  Mismatches   21;  Indels    4;  Gaps    3;

Qy       5 gacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtc 64
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 gacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagacccagtc 61

Qy      65 ctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccag 124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 ctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtgccag 121

Qy     125 caggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagaccgcca 184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     122 caggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacctcag 181

Qy     185 ccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgcca 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 agggccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctctgcca 241

Qy     245 ccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgccagt 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 ccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgccatg 301

Qy     305 ggcct--cagctgcccacacctcttcccggcccctgg-agttggcactgcagcagacagc 361
           | |||       |||||||:|:|||||||:|||||| |:||||||:|||||||||||||:
Db     302 gcccttcagctggcccacacytyttcccggscccctggaakttggcaytgcagcagacagy 361

Qy     362 tccctgggcaccaggcagctaaca-gacacagccgccagcccaaacagcagcggcatggg 420
           |||:||||||||||:|||:|||||| |||||||| ||||||||||||||||||||    ||
Db     362 tccytgggcaccagrcagytaacaggacacagcngccagcccaaacagcagcgggnatgg 421

Qy     421 cagc 424
             ||
Db     422 gggc 425


RESULT  11
AAC83263
ID   AAC83263 standard; DNA; 1751 BP.
XX
AC   AAC83263;
XX
DT   16-MAR-2001  (first entry)
XX
DE   GSEF promoter and partial coding sequence.
XX
KW   Transcription factor; gland-specific Ets transcription factor; GSEF;
KW   metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW   leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3; ds.
XX
OS   Homo sapiens.
XX
PN   WO200070092-A1.
XX
PD   23-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US13173.
XX
PR   14-MAY-1999;   99US-0134112.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kaufmann J,  Xin H,  Harrowe G;
XX
DR   WPI; 2001-041019/05.
XX
PT   Detecting metastatic and potential metastatic cancerous cells, useful
PT   for diagnosing, prognosing, grading and staging of cancers by detecting
PT   gland-specific Ets transcription factor gene product in a biological
PT   sample from a cell  -
XX
PS   Example 8; Page 88; 95pp; English.
XX
CC   This invention relates to a method for the detection and determination of
CC   the metastatic potential of a cell. The method comprises detecting a
CC   gland-specific Ets transcription factor (GSEF) gene product in a test
CC   sample. Detection of a GSEF gene product in the test sample in amount
CC   lower than that in a normal cell, is indicative of a cell with high
CC   metastatic potential. The method is useful for determining the metastatic
CC   potential of a cell, for the diagnosis and prognosis of cancer as well as
CC   grading and staging of cancers by detecting GSEF expression in a
CC   biological test sample. The method may also be used to monitor patients
CC   having a predisposition to develop a particular cancer. GSEF polypeptides
CC   are useful for producing antibodies, in cancer diagnosis, prognosis,
CC   grading, staging and management of breast and prostate tumours, and in
CC   detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC   useful in gene therapy. GSEF gene product expression levels can be used
CC   in conjunction with any tissue in which an alteration in GSEF gene
CC   product expression levels is associated with development of a
CC   cancer-associated phenotype. Cancers, which can be monitored include
CC   cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC   adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC   lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC   human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC   represents the the human GSEF promoter and partial coding sequence.
XX
SQ   Sequence 1751 BP; 369 A; 521 C; 497 G; 364 T; 0 other;


  Query Match             17.4%;  Score 329.8;  DB 22;  Length 1751;
  Best Local Similarity   97.7%;  Pred. No. 3.8e-60;
  Matches  345;  Conservative    0;  Mismatches    7;  Indels    1;  Gaps    1;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1400 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 1459

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1460 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 1519

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1520 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 1579

Qy     181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
           |     |||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    1580 tcagagggccacccccttgagggtggccaggcccccagt-gccaacctgagtgctgcctct 1638

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1639 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 1698

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcag 353
```

```
          |||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1699 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcag 1751


RESULT 12
AAA06617
ID   AAA06617 standard; cDNA; 278 BP.
XX
AC   AAA06617;
XX
DT   13-JUN-2000 (first entry)
XX
DE   Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:398.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
XX
PD   27-JAN-2000.
XX
PF   14-JUL-1999;   99WO-US15838.
XX
PR   14-JUL-1998;   98US-0115453.
PR   14-JUL-1998;   98US-0116134.
PR   23-SEP-1998;   98US-0159812.
PR   23-SEP-1998;   98US-0159822.
PR   15-JAN-1999;   99US-0232149.
PR   15-JAN-1999;   99US-0232880.
PR   09-APR-1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Harlocker SL,  Yuqiu J,  Xu J,  Mitcham JL;
XX
DR   WPI; 2000-171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 50; Page 239; 263pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82000 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 278 BP; 56 A; 85 C; 87 G; 49 T; 1 other;


  Query Match             13.9%;  Score 263;  DB 21;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 3e-46;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 254

Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 ttangtggctcaacaaggagaagg 278


RESULT 13
AAS63826
ID   AAS63826 standard; cDNA; 278 BP.
XX
AC   AAS63826;
XX
DT   29-JAN-2002 (first entry)
XX
DE   Human cDNA encodingEts transcription factor PDEF.
XX
KW   Human; prostate cancer; ss; cytostatic; immunostimulant; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200173032-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09919.
XX
PR   27-MAR-2000; 2000US-0536857.
PR   09-MAY-2000; 2000US-0568100.
PR   12-MAY-2000; 2000US-0570737.
PR   13-JUN-2000; 2000US-0593793.
PR   27-JUN-2000; 2000US-0605783.
PR   10-AUG-2000; 2000US-0636215.
PR   29-AUG-2000; 2000US-0651236.
PR   06-SEP-2000; 2000US-0657279.
PR   02-OCT-2000; 2000US-0679426.
PR   10-OCT-2000; 2000US-0685166.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Kalos MD;
PI   Fanger GR,  Retter MW,  Stolk JA,  Day CH,  Vedvick TS,  Carter D;
PI   Li SX,  Wang A,  Skeiky YAW,  Hepler WT,  Henderson RA;
XX
DR   WPI; 2001-639232/73.
XX
PT   New human prostate-specific polypeptides and polynucleotides useful for
PT   the diagnosis and treatment of cancer, especially prostate cancer -
XX
PS   Claim 1; Page 364-365; 579pp; English.
XX
CC   The invention relates to isolated prostate-specific
CC   polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC   antibodies raised against the polypeptides (or antigenic epitopes
CC   derived from them) and antigen-presenting cells expressing the
CC   polypeptides. The antibodies are useful for detecting the presence of
CC   cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC   the antigen-presenting cells are useful for stimulating and/or expanding
CC   T cells specific for a tumour protein, and for inhibiting the development
CC   of cancer especially prostate cancer. Compositions comprising the
CC   polynucleotide and/or polypeptide are useful for stimulating an immune
CC   response, and for treating cancer. The oligonucleotide is useful for
CC   detecting cancer. The present sequence is a prostate specific
CC   polynucleotide of the invention.
XX
SQ   Sequence 278 BP; 56 A; 85 C; 87 G; 49 T; 1 other;
```

```
  Query Match             13.9%;  Score 263;  DB 22;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 3e-46;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 254

Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 ttangtggctcaacaaggagaagg 278


RESULT  14
AAH93733
ID   AAH93733 standard; cDNA; 278 BP.
XX
AC   AAH93733;
XX
DT   04-OCT-2001  (first entry)
XX
DE   Human prostate-specific cDNA sequence Ets transcription factor.
XX
KW   Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW   cytostatic; gene therapy; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200151633-A2.
XX
PD   19-JUL-2001.
XX
PF   16-JAN-2001; 2001WO-US01574.
XX
PR   14-JAN-2000; 2000US-0483672.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Fanger GR,  Day CH,  Retter MW,  Stolk JA,  Skeiky YAW;
PI   Wang A,  Meagher MJ;
XX
DR   WPI; 2001-425873/45.
XX
PT   New polynucleotide encoding a prostate-specific protein, for
PT   diagnosing, monitoring and treating prostate cancer in a patient and
PT   for use in vaccines -
XX
PS   Claim 1; Page 363; 543pp; English.
XX
CC   The present invention describes polynucleotide sequences (I) which encode
CC   prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC   and can be used in vaccine production and gene therapy. (I), (II),
CC   antibodies to (II), fusion proteins comprising (II), and isolated
CC   T cells prepared using (I) or (II) are used treat cancer in a patient.
CC   (I) and the antibodies are also used in the detection of cancer in a
CC   patient. The cancer that is diagnosed or treated is particularly
CC   prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC   (I) can be used for monitoring the progression of cancer in a patient.
CC   (I) and (II) can also be used to improve diagnostic and therapeutic
CC   methods for prostate cancer. They can indicate the level of metastasis
CC   as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC   AAM01318 represent polynucleotide and amino acid sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 278 BP; 56 A; 85 C; 87 G; 49 T; 1 other;


  Query Match             13.9%;  Score 263;  DB 22;  Length 278;
  Best Local Similarity   99.6%;  Pred. No. 3e-46;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 254

Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 ttangtggctcaacaaggagaagg 278


RESULT  15
AAH85047
ID   AAH85047 standard; cDNA; 278 BP.
XX
AC   AAH85047;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human prostate-specific cDNA sequence Ets transcription factor PDEF.
XX
KW   Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW   chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW   prostate specific antigen; PSA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200134802-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30904.
XX
PR   12-NOV-1999;   99US-0439313.
PR   18-NOV-1999;   99US-0443686.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Dillon DC,  Mitcham JL,  Harlocker SL,  Jiang Y,  Reed SG;
PI   Kalos MD,  Retter MW,  Stolk JA,  Day CH,  Skeiky YAW,  Wang A;
XX
DR   WPI; 2001-308785/32.
XX
PT   Isolated polypeptide comprising at least an immunogenic portion of a
PT   prostate-specific protein, useful in the diagnosis and therapy of
PT   prostate cancer -
XX
PS   Claim 31; Page 263; 325pp; English.
XX
CC   The present invention describes an isolated polypeptide (P1) comprising
CC   at least an immunogenic portion of a prostate-specific protein, or its
```

```
CC   variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
CC   (N1) have cytostatic activity and can be used in vaccine production.
CC   The polypeptides, nucleic acids and antibodies from the present
CC   invention are useful in the diagnosis and therapy of prostate cancer.
CC   Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC   in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC   region. Prostate specific antigen (PSA) P501S was located on
CC   chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC   polynucleotide and polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 278 BP; 56 A; 85 C; 87 G; 49 T; 1 other;


  Query Match            13.9%;  Score 263;  DB 22;  Length 278;
  Best Local Similarity  99.6%;  Pred. No. 3e-46;
  Matches  263;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    981 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 agcagttccgccagcgctcgcccctgggtggggatgtgctgcacgcccacctggacatct 74

Qy   1041 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 ggaagtcagcggcctggatgaaagagcggacttcacctggggcgattcactactgtgcct 134

Qy   1101 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 cgaccagtgaggagagctggaccgacagcgaggtggactcatcatgctccgggcagccca 194

Qy   1161 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 tccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttca 254

Qy   1221 ttaggtggctcaacaaggagaagg 1244
          ||| ||||||||||||||||||||
Db    255 ttangtggctcaacaaggagaagg 278


Search completed: August  6, 2002, 21:01:06
Job time: 7725 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 18:17:31 ; Search time 3380.83 Seconds
(without alignments)
12120.248 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US100_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US101_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1894	100.0	1894	1	PCT-US01-03640-1	Sequence 1, Appli
	2	1894	100.0	1894	32	US-09-841-963A-1	Sequence 1, Appli
	3	1870	98.7	1894	1	PCT-US99-17470-1	Sequence 1, Appli
	4	1870	98.7	1894	15	US-09-126-945-1	Sequence 1, Appli
	5	1870	98.7	1894	15	US-09-126-945A-1	Sequence 1, Appli
	6	1870	98.7	1894	15	US-09-126-945B-1	Sequence 1, Appli
	7	1870	98.7	1894	22	US-09-570-593-1	Sequence 1, Appli
	8	1868.4	98.6	1902	14	US-09-064-645-14	Sequence 14, Appl
	9	1868.4	98.6	1902	14	US-09-064-645-15	Sequence 15, Appl
	10	1868.4	98.6	1902	22	US-09-567-970-14	Sequence 14, Appl
	11	1868.4	98.6	1902	22	US-09-567-970-15	Sequence 15, Appl
	12	1868.4	98.6	1918	28	US-09-718-159-1	Sequence 1, Appli
	13	1865.2	98.5	3317	22	US-09-570-593-12	Sequence 12, Appl
	14	1861.4	98.3	1905	33	US-09-866-356-2	Sequence 2, Appli
	15	1837.8	97.0	1907	59	US-60-209-009-74	Sequence 74, Appl
	16	1824.4	96.3	1916	56	US-60-172-373-4734	Sequence 4734, Ap
	17	1803	95.2	1841	12	US-08-845-995-12	Sequence 12, Appl
	18	1019.8	53.8	1087	1	PCT-US00-05881-215	Sequence 215, App
	19	1019.8	53.8	1087	34	US-09-925-298-215	Sequence 215, App
c	20	790.6	41.7	1188	59	US-60-209-009-75	Sequence 75, Appl
c	21	790.6	41.7	1188	71	US-60-324-185-16637	Sequence 16637, A
	22	755.2	39.9	1701	75	US-60-360-207-11811	Sequence 11811, A
	23	626.4	33.1	1704	1	PCT-US01-03640-3	Sequence 3, Appli
	24	626.4	33.1	1704	32	US-09-841-963A-3	Sequence 3, Appli
	25	484.8	25.6	792	14	US-09-016-434-840	Sequence 840, App
	26	459.4	24.3	478	25	US-09-644-869-2777	Sequence 2777, Ap
	27	459.4	24.3	478	25	US-09-649-163-2650	Sequence 2650, Ap
	28	434.6	22.9	609	20	US-09-535-897-20568	Sequence 20568, A
	29	425	22.4	450	19	US-09-528-409-87128	Sequence 87128, A
	30	425	22.4	450	35	US-09-933-524-87128	Sequence 87128, A
	31	425	22.4	450	35	US-09-933-524A-87128	Sequence 87128, A

32	418	22.1	442	19	US-09-528-409-84444	Sequence 84444, A
33	418	22.1	442	35	US-09-933-524-84444	Sequence 84444, A
34	418	22.1	442	35	US-09-933-524A-84444	Sequence 84444, A
35	416.6	22.0	457	16	US-09-240-371-5588	Sequence 5588, Ap
36	416.6	22.0	457	16	US-09-277-227-20468	Sequence 20468, A
37	416.6	22.0	457	34	US-09-909-627-20468	Sequence 20468, A
38	416.6	22.0	457	34	US-09-919-724-5588	Sequence 5588, Ap
39	416.2	22.0	453	19	US-09-528-409-84270	Sequence 84270, A
40	416.2	22.0	453	35	US-09-933-524-84270	Sequence 84270, A
41	416.2	22.0	453	35	US-09-933-524A-84270	Sequence 84270, A
42	415.6	21.9	463	16	US-09-293-972-7218	Sequence 7218, Ap
43	415.6	21.9	463	34	US-09-904-939-7218	Sequence 7218, Ap
44	414.4	21.9	437	19	US-09-528-409-68832	Sequence 68832, A
45	414.4	21.9	437	35	US-09-933-524-68832	Sequence 68832, A

ALIGNMENTS

```
RESULT   1
PCT-US01-03640-1
; Sequence 1, Application PC/TUS0103640
; GENERAL INFORMATION:
;  APPLICANT: MUSC Foundation for Research  Development
;  TITLE OF INVENTION: Methods and compositions for the diagnosis and treatment of cance
;  TITLE OF INVENTION:  based on ESF
;  FILE REFERENCE: 9175-015-228
;  CURRENT APPLICATION NUMBER: PCT/US01/03640
;  CURRENT FILING DATE:  2001-02-09
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (416)..(1423)
PCT-US01-03640-1


  Query Match           100.0%;  Score 1894;  DB 1;  Length 1894;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1894;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
Db       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
Db     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
Db     181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
Db     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
Db     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
Db     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
Db     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
Db     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
Db     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
Db     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
Db     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
Db     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
Db     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
Db     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
Db     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
Db     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
Db    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
Db    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
Db    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
Db    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
Db    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
Db    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
Db    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
Db    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
```

```
Db   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

```
RESULT   2
US-09-841-963A-1
; Sequence 1, Application US/09841963A
; GENERAL INFORMATION:
;  APPLICANT: Watson, Dennis K.
;  APPLICANT:  Papas, Takis S. (Deceased)
;  APPLICANT:  Papas, Tula C. (Legal Representative)
;  TITLE OF INVENTION: Methods and compositions for the diagnosis and treatment of cance
;  TITLE OF INVENTION:  based on transcription factor ETS2
;  FILE REFERENCE: 10545-015-999
;  CURRENT APPLICATION NUMBER: US/09/841,963A
;  CURRENT FILING DATE:  2001-04-25
;  PRIOR APPLICATION NUMBER: PCT/US99/27805
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: 06/109,850
;  PRIOR FILING DATE: 1998-11-25
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (416)..(1423)
US-09-841-963A-1
```

```
  Query Match           100.0%;  Score 1894;  DB 32;  Length 1894;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1894;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
Db       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180
Db     121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccacccttgagggtggccaggccccagtggccaacctgagtgctgcctct 240
Db     181 gccaccagccacccttgagggtggccaggccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
Db     241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
Db     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
Db     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
Db     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
Db     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
Db     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
Db     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
Db     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
Db     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
Db     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
Db     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
Db     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
Db     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
Db    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
Db    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
Db    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
Db    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
```

```
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          ||||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

```
RESULT   3
PCT-US99-17470-1
; Sequence 1, Application PC/TUS9917470
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  APPLICANT: Beth Israel Deaconess Medical Center
;  APPLICANT: Libermann, Towia A.
;  APPLICANT: Oettgen, Joerg P.
;  APPLICANT: Kunsch, Charles A.
;  APPLICANT: Endress, Gregory A.
;  APPLICANT: Rosen, Craig A.
;  TITLE OF INVENTION: Prostate Derived Ets Factor
;  FILE REFERENCE: 1488.109PC01
;  CURRENT APPLICATION NUMBER: PCT/US99/17470
;  CURRENT FILING DATE: 1999-08-02
;  EARLIER APPLICATION NUMBER: US 09/126,945
;  EARLIER FILING DATE: 1998-07-31
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
PCT-US99-17470-1
```

Query Match 98.7%; Score 1870; DB 1; Length 1894;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1879; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy    421 cagcgccagcccgggtctgagcagcgtatccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagcgccagcccgggtctgagcagcgtatccccagccacctcctgctgccccccgacac 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
```

```
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894


RESULT   4
US-09-126-945-1
; Sequence 1, Application US/09126945
; GENERAL INFORMATION:
;  APPLICANT: Libermann, Towia A.
;  APPLICANT: Oettgen, Joerg P.
;  APPLICANT: Kunsch, Chuck
;  APPLICANT: Endress, Gregory A.
;  APPLICANT: Rosen, Craig A.
;  TITLE OF INVENTION: Prostate Derived Ets Factor
;  FILE REFERENCE: 1488.1090000
;  CURRENT APPLICATION NUMBER: US/09/126,945
;  CURRENT FILING DATE: 1998-07-31
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-126-945-1


  Query Match            98.7%;  Score 1870;  DB 15;  Length 1894;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
```

```
Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894


RESULT   5
US-09-126-945A-1
; Sequence 1, Application US/09126945A
; GENERAL INFORMATION:
;  APPLICANT:  Libermann, Towia A.
;  APPLICANT:  Oettgen, Joerg P.
;  APPLICANT:  Kunsch, Charles A.
;  APPLICANT:  Endress, Gregory A.
;  APPLICANT:  Rosen, Craig A.
;  TITLE OF INVENTION:  Prostate Derived Ets Factor
;  FILE REFERENCE:  1488.1090000
;  CURRENT APPLICATION NUMBER: US/09/126,945A
;  CURRENT FILING DATE:  1998-07-31
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-126-945A-1


  Query Match             98.7%;  Score 1870;  DB 15;  Length 1894;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgacac 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
```

```
Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

```
RESULT   6
US-09-126-945B-1
; Sequence 1, Application US/09126945B
; GENERAL INFORMATION:
;  APPLICANT:  Libermann, Towia A.
;  APPLICANT:  Oettgen, Joerg P.
;  APPLICANT:  Kunsch, Charles A.
;  APPLICANT:  Endress, Gregory A.
;  APPLICANT:  Rosen, Craig A.
;  TITLE OF INVENTION:  Prostate Derived Ets Factor
;  FILE REFERENCE:  1488.1090000
;  CURRENT APPLICATION NUMBER: US/09/126,945B
;  CURRENT FILING DATE:  1998-07-31
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-126-945B-1


  Query Match             98.7%;  Score 1870;  DB 15;  Length 1894;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy    121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ccagcaggcagccctgggctggggtaggggactccctacaggcacgcagccctgagacc 180

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
```

```
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894


RESULT   7
US-09-570-593-1
; Sequence 1, Application US/09570593
; GENERAL INFORMATION:
;  APPLICANT:  Kaufmann, Joerg
;  APPLICANT:  Xin, Hong
;  APPLICANT:  Harrowe, Greg
;  TITLE OF INVENTION:  EXPRESSION OF ETS-DOMAIN PROTEINS IN
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE:  2300-1556
;  CURRENT APPLICATION NUMBER: US/09/570,593
;  CURRENT FILING DATE:  2000-05-12
;  PRIOR APPLICATION NUMBER:  60/134,112
;  PRIOR FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (416)...(1423)
;   OTHER INFORMATION: Human JKETS.
US-09-570-593-1


  Query Match             98.7%;  Score 1870;  DB 22;  Length 1894;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1879;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180

Qy     181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
           |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240

Qy     241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     301 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 360

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
```

```
Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 600

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db    661 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 720

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1320

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1321 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1860

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1861 cccgggaatggataataaagatactagagaactg 1894
```

RESULT 8
US-09-064-645-14
; Sequence 14, Application US/09064645
; GENERAL INFORMATION:
; APPLICANT: PATRICIA A. BILLING-MEDEL
; APPLICANT: MAURICE COHEN
; APPLICANT: TRACEY L. COLPITTS
; APPLICANT: PAULA N. FRIEDMAN
; APPLICANT: MICHAEL R. KLASS
; APPLICANT: LISA ROBERTS-RAPP
; APPLICANT: JOHN C. RUSSELL
; APPLICANT: STEPHEN D. STROUPE
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,645
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,995
; FILING DATE: 25-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6085.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1902 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-064-645-14

```
  Query Match             98.6%;  Score 1868.4;  DB 14;  Length 1902;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1878;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
Db       8 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 67

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
Db      68 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 127

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
Db     128 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 187

Qy     181 gccaccagccacccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
Db     188 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 247

Qy     241 gccaccagccctgctggcccctggttccgctggcccccccagatgcctggctgagacacgc 300
Db     248 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 307

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
Db     308 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 367

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
Db     368 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 427

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
Db     428 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 487

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
Db     488 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 547

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
Db     548 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 607

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
Db     608 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 667

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
Db     668 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 727

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
Db     728 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 787

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
Db     788 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 847

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
Db     848 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 907

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
Db     908 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 967

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
Db     968 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1027

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
Db    1028 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1087

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
Db    1088 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1147

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
Db    1148 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1207

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
Db    1208 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1267

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
Db    1268 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1327

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
Db    1328 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1387

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
Db    1388 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1447

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
Db    1448 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1507

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
Db    1508 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1567

Qy    1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
Db    1568 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGG 1627

Qy    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
Db    1628 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1687

Qy    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
Db    1688 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1747

Qy    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
Db    1748 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1807

Qy    1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
Db    1808 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1867

Qy    1861 cccgggaatggataataaagatactagagaactg 1894
Db    1868 CCCGGGAATGGATAATAAAGATACTAGAGAACTG 1901


RESULT   9
US-09-064-645-15
; Sequence 15, Application US/09064645
;  GENERAL INFORMATION:
;    APPLICANT:  PATRICIA A. BILLING-MEDEL
;    APPLICANT:  MAURICE COHEN
;    APPLICANT:  TRACEY L. COLPITTS
;    APPLICANT:  PAULA N. FRIEDMAN
;    APPLICANT:  MICHAEL R. KLASS
;    APPLICANT:  LISA ROBERTS-RAPP
;    APPLICANT:  JOHN C. RUSSELL
```

```
;    APPLICANT:  STEPHEN D. STROUPE
;    TITLE OF INVENTION:  REAGENTS AND METHODS USEFUL
;    TITLE OF INVENTION:  FOR DETECTING DISEASES OF THE PROSTATE
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/064,645
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/845,995
;      FILING DATE:  25-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6085.US.P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1902 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-064-645-15


  Query Match             98.6%;  Score 1868.4;  DB 14;  Length 1902;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1878;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 67

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 127

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 187

Qy    181 gccaccagccacccctttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
            |    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCCAGTGGCCAACCTGAGTGCTGCCTCT 247

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 307

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    308 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 367

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 427

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 487

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 547

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    548 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 607

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 667

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    668 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 727

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 787

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 847

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 907

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 967

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1027

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1087

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1147

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1207

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1267

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1327

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1328 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1387

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1447

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1448 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1507

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1508 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1567

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1568 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGG 1627

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1628 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1687

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1688 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1747

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1748 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1807

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1808 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1867

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1868 CCCGGGAATGGATAATAAAGATACTAGAGAACTG 1901


RESULT 10
US-09-567-970-14
; Sequence 14, Application US/09567970
; GENERAL INFORMATION:
;  APPLICANT:  Abbott Laboratories
;  APPLICANT:  Billing-Medel, Patricia A.
;  APPLICANT:  Colpitts, Tracey L.
;  APPLICANT:  Friedman, Paula N.
;  APPLICANT:  Klass, Michael R.
;  APPLICANT:  Roberts-Rapp, Lisa
;  APPLICANT:  Russell, John C.
;  APPLICANT:  Stroupe, Stephen D.
;  TITLE OF INVENTION:  REAGENTS AND METHODS USEFUL FOR
;  TITLE OF INVENTION:  DETECTING DISEASES OF THE PROSTATE
;  FILE REFERENCE:  6085.US.P2
;  CURRENT APPLICATION NUMBER: US/09/567,970
;  CURRENT FILING DATE:  2000-05-10
;  PRIOR APPLICATION NUMBER:  US 09/064,645
;  PRIOR FILING DATE:  1998-04-22
;  PRIOR APPLICATION NUMBER:  US 08/845,995
;  PRIOR FILING DATE:  1997-04-25
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 1902
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-567-970-14


  Query Match             98.6%;  Score 1868.4;  DB 22;  Length 1902;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1878;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      8 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 67

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     68 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 127

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 187

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 247

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 307

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    308 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 367

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    368 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 427

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    428 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 487

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    488 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 547

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    548 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 607

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    608 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 667

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    668 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 727

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    728 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 787

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    788 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 847

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    848 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 907

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    908 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 967

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    968 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1027

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1087

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1147

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1207

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1267

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
```

```
Db   1268 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1327

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1328 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1387

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1447

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1448 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1507

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1508 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1567

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1568 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcaggg 1627

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1628 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1687

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1688 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1747

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1748 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1807

Qy   1801 ccacacctggcatggtgcagggagacatctgcaccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1808 ccacacctggcatggtgcagggagacatctgcaccctgagttgggcagccaggagtgcc 1867

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||||
Db   1868 cccgggaatggataataaagatactagagaactg 1901


RESULT 11
US-09-567-970-15
; Sequence 15, Application US/09567970
; GENERAL INFORMATION:
;  APPLICANT:  Abbott Laboratories
;  APPLICANT:  Billing-Medel, Patricia A.
;  APPLICANT:  Colpitts, Tracey L.
;  APPLICANT:  Friedman, Paula N.
;  APPLICANT:  Klass, Michael R.
;  APPLICANT:  Roberts-Rapp, Lisa
;  APPLICANT:  Russell, John C.
;  APPLICANT:  Stroupe, Stephen D.
;  TITLE OF INVENTION:  REAGENTS AND METHODS USEFUL FOR
;  TITLE OF INVENTION:  DETECTING DISEASES OF THE PROSTATE
;  FILE REFERENCE:  6085.US.P2
;  CURRENT APPLICATION NUMBER: US/09/567,970
;  CURRENT FILING DATE:  2000-05-10
;  PRIOR APPLICATION NUMBER:  US 09/064,645
;  PRIOR FILING DATE:  1998-04-22
;  PRIOR APPLICATION NUMBER:  US 08/845,995
;  PRIOR FILING DATE:  1997-04-25
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 1902
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-567-970-15
```

```
  Query Match             98.6%;  Score 1868.4;  DB 22;  Length 1902;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1878;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 67

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      68 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 127

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 187

Qy     181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
             |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     188 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 247

Qy     241 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     248 gccaccagccctgctggcccctggttccgctggccccccagatgcctggctgagacacgc 307

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     308 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 367

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     368 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 427

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     428 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 487

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     488 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 547

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     548 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 607

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     608 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 667

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db     668 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 727

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     728 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 787

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     788 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 847

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     848 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 907

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     908 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 967

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     968 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1027
```

```
Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1087

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1147

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1207

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1267

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1327

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   1328 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1387

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1447

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1448 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1507

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1508 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1567

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1568 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcaggg 1627

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1628 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1687

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1688 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1747

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1748 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1807

Qy   1801 ccacacctggcatggtgcagggagacatctgcaccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1808 ccacacctggcatggtgcagggagacatctgcaccctgagttgggcagccaggagtgcc 1867

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          |||||||||||||||||||||||||||||||
Db   1868 cccgggaatggataataaagatactagagaactg 1901
```

RESULT 12
US-09-718-159-1
; Sequence 1, Application US/09718159
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Lin, Richard
; APPLICANT: Luke, May
; APPLICANT: Monteclaro, Felipe
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heuit, Pamela Toy
; APPLICANT: Xuan, Jian-ai
; TITLE OF INVENTION: DNA Encoding a Novel ETS Polypeptide
; FILE REFERENCE: 51790AUSM1
; CURRENT APPLICATION NUMBER: US/09/718,159
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/168,182
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1918
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (422)..(1429)
US-09-718-159-1

Query Match 98.6%; Score 1868.4; DB 28; Length 1918;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      7 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 66

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 126

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 186

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
          |     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 246

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 306

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    307 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 366

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 426

Qy    421 cagcgccagcccgggtctgagcagcgtatccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 cagcgccagcccgggtctgagcagcgtatccccagccacctcctgctgccccccgacac 486

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 546

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    547 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 606

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    607 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 666

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    667 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 726

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    727 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 786

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    787 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 846

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    847 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 906

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    907 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 966

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    967 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1026

Qy   1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1027 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1086

Qy   1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1146

Qy   1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1206

Qy   1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1266

Qy   1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1267 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 1326

Qy   1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1327 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1386

Qy   1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1387 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1446

Qy   1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1447 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1506

Qy   1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   1507 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 1566

Qy   1561 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1567 gcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcaggg 1626

Qy   1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1627 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1686

Qy   1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1687 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1746

Qy   1741 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1747 acaaaggccacaggcagtccaggcctctctctgctccatcccccctgcctcccattctgca 1806

Qy   1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1807 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 1866

Qy   1861 cccgggaatggataataaagatactagagaactg 1894
          ||||||||||||||||||||||||||||||||||
Db   1867 cccgggaatggataataaagatactagagaactg 1900


RESULT 13
US-09-570-593-12
; Sequence 12, Application US/09570593
; GENERAL INFORMATION:
;  APPLICANT:  Kaufmann, Joerg
;  APPLICANT:  Xin, Hong
;  APPLICANT:  Harrowe, Greg
;  TITLE OF INVENTION:  EXPRESSION OF ETS-DOMAIN PROTEINS IN
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE:  2300-1556
;  CURRENT APPLICATION NUMBER: US/09/570,593
;  CURRENT FILING DATE:  2000-05-12
;  PRIOR APPLICATION NUMBER:  60/134,112
;  PRIOR FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 12
;   LENGTH: 3317
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)...(3317)
;   OTHER INFORMATION: GSEF-encoding sequence with promoter (Figs. 2A-2B)
;   NAME/KEY: promoter
;   LOCATION: (1)...(1380)
;   OTHER INFORMATION: Promoter
;   NAME/KEY: CDS
;   LOCATION: (1815)...(2819)
;   OTHER INFORMATION: GSEF coding sequence
US-09-570-593-12


  Query Match             98.5%;  Score 1865.2;  DB 22;  Length 3317;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches 1876;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1400 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 1459

Qy     61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1460 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 1519

Qy    121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1520 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 1579

Qy    181 gccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 240
            |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1580 tcagagggccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcctct 1639

Qy    241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1640 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 1699

Qy    301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1700 cagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagacag 1759

Qy    361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1760 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 1819

Qy    421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1820 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 1879

Qy    481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1880 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 1939

Qy    541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   1940 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctccta 1999

Qy    601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2000 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 2059

Qy    661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db   2060 tcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagccccagc 2119

Qy    721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2120 gggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 2179

Qy    781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2180 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 2239

Qy    841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2240 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 2299

Qy    901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2300 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 2359

Qy    961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtgct 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   2360 gctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggagatgtgct 2419

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2420 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 2479

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2480 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 2539

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2540 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 2599

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2600 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 2659

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2660 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaacta 2719

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   2720 cgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagccaga 2779

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2780 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 2839

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2840 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 2899

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2900 ggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggaggg 2959

Qy    1561 gcaaccaactgccccaggggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2960 gcaaccaactgccccaggggggatatgggtcctctggggccttcgggaccctggggcaggg 3019

Qy    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3020 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 3079

Qy    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3080 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 3139

Qy    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3140 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 3199

Qy    1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   3200 ccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtgcc 3259

Qy    1861 cccgggaatggataataaagatactagagaactg 1894
           |||||||||||||||| ||||||||||||||||
Db   3260 cccgggaatggataatagagatactagagaactg 3293
```

```
RESULT  14
US-09-866-356-2
; Sequence 2, Application US/09866356
;   GENERAL INFORMATION:
;        APPLICANT: Bandman, Olga
;                   Corley, Neil C.
;                   Guegler, Karl J.
;                   Lal, Preeti
;        TITLE OF INVENTION: PROSTATE-ASSOCIATED ETS PROTEIN
;        NUMBER OF SEQUENCES: 4
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/866,356
;             FILING DATE: 29-May-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/055,113
;             FILING DATE: 1998-04-03
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0501 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1905 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: PROSTUT12
;             CLONE: 1813005
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-866-356-2
```

```
  Query Match             98.3%;  Score 1861.4;  DB 33;  Length 1905;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1871;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15 GTCTGACTTCCTCCCAGCACATTCCTGCACTCTGCCGTGTCCACACTGCCCCACAGACCC 74

Qy      61 agtcctccaagcctgctgccagctccctgcaagcccctcaggttgggccttgccacggtg 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 AGTCCTCCAAGCCTGCTGCCAGCTCCCTGCAAGCCCCTCAGGTTGGGCCTTGCCACGGTG 134

Qy     121 ccagcaggcagccctgggctgggggtaggggactccctacaggcacgcagccctgagacc 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 CCAGCAGGCAGCCCTGGGCTGGGGGTAGGGGACTCCCTACAGGCACGCAGCCCTGAGACC 194

Qy     181 gccaccagcccacccctgagggtggccaggccccagtggccaacctgagtgctgcctct 240
           |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 TCAGAGGGCCACCCCTTGAGGGTGGCCAGGCCCCAGTGGCCAACCTGAGTGCTGCCTCT 254

Qy     241 gccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacacgc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     255 GCCACCAGCCCTGCTGGCCCCTGGTTCCGCTGGCCCCCCAGATGCCTGGCTGAGACACGC 314

Qy     301 cagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagacag 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 CAGTGGCCTCAGCTGCCCACACCTCTTCCCGGCCCCTGAAGTTGGCACTGCAGCAGACAG 374

Qy     361 ctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     375 CTCCCTGGGCACCAGGCAGCTAACAGACACAGCCGCCAGCCCAAACAGCAGCGGCATGGG 434

Qy     421 cagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgccccccgacac 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     435 CAGCGCCAGCCCGGGTCTGAGCAGCGTATCCCCCAGCCACCTCCTGCTGCCCCCCGACAC 494

Qy     481 ggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacggga 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     495 GGTGTCGCGGACAGGCTTGGAGAAGGCGGCAGCGGGGGCAGTGGGTCTCGAGAGACGGGA 554

Qy     541 ctggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctccta 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     555 CTGGAGTCCCAGTCCACCCGCCACGCCCGAGCAGGGCCTGTCCGCCTTCTACCTCTCCTA 614

Qy     601 ctttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagcag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     615 CTTTGACATGCTGTACCCTGAGGACAGCAGCTGGGCAGCCAAGGCCCCTGGGGCCAGCAG 674

Qy     661 tcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagccccagc 720
           ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db     675 TCGGGAGGAGCCACCTGAGGAGCCTGAGCAGTGCCCGGTCATTGACAGCCAAGCCCCAGC 734

Qy     721 gggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcaggt 780
           |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     735 GGGCAGCCTGGACTTGGTGCCCGGCGGGCTGACCTTGGAGGAGCACTCGCTGGAGCAGGT 794

Qy     781 gcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctcaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     795 GCAGTCCATGGTGGTGGGCGAAGTGCTCAAGGACATCGAGACGGCCTGCAAGCTGCTCAA 854

Qy     841 catcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtggac 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     855 CATCACCGCAGATCCCATGGACTGGAGCCCCAGCAATGTGCAGAAGTGGCTCCTGTGGAC 914

Qy     901 agagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaagga 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     915 AGAGCACCAATACCGGCTGCCCCCCATGGGCAAGGCCTTCCAGGAGCTGGCGGGCAAGGA 974

Qy     961 gctgtgcgccatgtcggaggagcagttccgccagcgctcgccccctgggtggggatgtgct 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     975 GCTGTGCGCCATGTCGGAGGAGCAGTTCCGCCAGCGCTCGCCCCTGGGTGGGGATGTGCT 1034

Qy    1021 gcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacctgg 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1035 GCACGCCCACCTGGACATCTGGAAGTCAGCGGCCTGGATGAAAGAGCGGACTTCACCTGG 1094

Qy    1081 ggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggactc 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1095 GGCGATTCACTACTGTGCCTCGACCAGTGAGGAGAGCTGGACCGACAGCGAGGTGGACTC 1154

Qy    1141 atcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaagcc 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1155 ATCATGCTCCGGGCAGCCCATCCACCTGTGGCAGTTCCTCAAGGAGTTGCTACTCAAGCC 1214

Qy    1201 ccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattga 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1215 CCACAGCTATGGCCGCTTCATTAGGTGGCTCAACAAGGAGAAGGGCATCTTCAAAATTGA 1274

Qy    1261 ggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1275 GGACTCAGCCCAGGTGGCCCGGCTGTGGGGCATCCGCAAGAACCGTCCCGCCATGAACTA 1334

Qy    1321 cgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagccaga 1380
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1335 CGACAAGCTGAGCCGCTCCATCCGCCAGTATTACAAGAAGGGCATCATCCGGAAGCCAGA 1394

Qy    1381 catctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccagggcc 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1395 CATCTCCCAGCGCCTCGTCTACCAGTTCGTGCACCCCATCTGAGTGCCTGGCCCAGGGCC 1454

Qy    1441 tgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgagatg 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1455 TGAAACCCGCCCTCAGGGGCCTCTCTCCTGCCTGCCCTGCCTCAGCCAGGCCCTGAGATG 1514

Qy    1501 ggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggaggg 1560
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1515 GGGGAAAACGGGCAGTCTGCTCTGCTGCTCTGACCTTCCAGAGCCCAAGGTCAGGGAGGG 1574

Qy    1561 gcaaccaactgccccaggggggatatgggtcctctggggccttcgggaccatggggcaggg 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    1575 GCAACCAACTGCCCCAGGGGGATATGGGTCCTCTGGGGCCTTCGGGACCCTGGGGCAGGG 1634

Qy    1621 gtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctccc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1635 GTGCTTCCTCCTCAGGCCCAGCTGCTCCCCTGGAGGACAGAGGGAGACAGGGCTGCTCCC 1694

Qy    1681 caacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgactcg 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1695 CAACACCTGCCTCTGACCCCAGCATTTCCAGAGCAGAGCCTACAGAAGGGCAGTGACTCG 1754

Qy    1741 acaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctgca 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1755 ACAAAGGCCACAGGCAGTCCAGGCCTCTCTCTGCTCCATCCCCCTGCCTCCCATTCTGCA 1814

Qy    1801 ccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtgcc 1860
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    1815 CCACACCTGGCATGGTGCAGGGAGACATCTGCACCCCTGAGTTGGGCAGCCAGGAGTGCC 1874

Qy    1861 cccgggaatggataataaagatactag 1887
           |||||||||||||||||||||||||||
Db    1875 CCCGGGAATGGATAATAAAGATACTAG 1901


RESULT  15
US-60-209-009-74
; Sequence 74, Application US/60209009
; GENERAL INFORMATION:
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Pearson, Cecelia I.
;  TITLE OF INVENTION:      GENES EXPRESSED IN PROSTATE CANCER
;  FILE REFERENCE:  PA-0027 P
;  CURRENT APPLICATION NUMBER: US/60/209,009
;  CURRENT FILING DATE:  2000-06-01
;  NUMBER OF SEQ ID NOS:  501
```

```
;  SOFTWARE:  PERL Program
; SEQ ID NO 74
;   LENGTH: 1907
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 234340.7
US-60-209-009-74


  Query Match             97.0%;  Score 1837.8;  DB 59;  Length 1907;
  Best Local Similarity   99.0%;  Pred. No. 0;
  Matches 1870;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    2;

Qy      1 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 gtctgacttcctcccagcacattcctgcactctgccgtgtccacactgccccacagaccc 74

Qy     61 agtcctccaagcctgctgccagct-ccctgcaagcccctcaggttgggccttgccacggt 119
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     75 agtcctccaagcctgctgccagctgccctgcaagcccctcaggttgggccttgccacggt 134

Qy    120 gccagcaggcagccct-gggctgggggtaggggactccctacaggcacgcagccctgaga 178
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    135 gccagcaggcagccctggggctgggggtaggggactccctacaggcacgcagccctgaga 194

Qy    179 ccgccaccagccaccccttgagggtggccaggcccccagtggccaacctgagtgctgcct 238
          || |     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 cctcagagggccaccccttgagggtggccaggccccagtggccaacctgagtgctgcct 254

Qy    239 ctgccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacac 298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 ctgccaccagccctgctggcccctggttccgctggcccccagatgcctggctgagacac 314

Qy    299 gccagtggcctcagctgcccacacctcttcccggcccctggagttggcactgcagcagac 358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 gccagtggcctcagctgcccacacctcttcccggcccctgaagttggcactgcagcagac 374

Qy    359 agctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatg 418
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 agctccctgggcaccaggcagctaacagacacagccgccagcccaaacagcagcggcatg 434

Qy    419 ggcagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgac 478
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 ggcagcgccagcccgggtctgagcagcgtatcccccagccacctcctgctgcccccgac 494

Qy    479 acggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgg 538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 acggtgtcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagagacgg 554

Qy    539 gactggagtcccagtccacccgccacgcccgagcagggcctgtccgccttcgacctctcc 598
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    555 gactggagtcccagtccacccgccacgcccgagcagggcctgtccgccttctacctctcc 614

Qy    599 tactttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagc 658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 tactttgacatgctgtaccctgaggacagcagctgggcagccaaggcccctggggccagc 674

Qy    659 agtcgggaggagccacctgaccagcctgagcagtgcccggtcattgacagccaagcccca 718
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    675 agtcgggaggagccacctgaggagcctgagcagtgcccggtcattgacagccaagcccca 734

Qy    719 gcgggcaccctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcag 778
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 gcgggcagcctggacttggtgcccggcgggctgaccttggaggagcactcgctggagcag 794

Qy    779 gtgcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctc 838
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 gtgcagtccatggtggtgggcgaagtgctcaaggacatcgagacggcctgcaagctgctc 854

Qy    839 aacatcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtgg 898
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    855 aacatcaccgcagatcccatggactggagccccagcaatgtgcagaagtggctcctgtgg 914

Qy    899 acagagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaag 958
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 acagagcaccaataccggctgccccccatgggcaaggccttccaggagctggcgggcaag 974

Qy    959 gagctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtg 1018
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 gagctgtgcgccatgtcggaggagcagttccgccagcgctcgcccctgggtggggatgtg 1034

Qy   1019 ctgcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacct 1078
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 ctgcacgcccacctggacatctggaagtcagcggcctggatgaaagagcggacttcacct 1094

Qy   1079 ggggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggac 1138
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1095 ggggcgattcactactgtgcctcgaccagtgaggagagctggaccgacagcgaggtggac 1154

Qy   1139 tcatcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaag 1198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1155 tcatcatgctccgggcagcccatccacctgtggcagttcctcaaggagttgctactcaag 1214

Qy   1199 ccccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaatt 1258
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1215 ccccacagctatggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaatt 1274

Qy   1259 gaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaac 1318
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1275 gaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaac 1334

Qy   1319 tgcgacaagctgagccgctccatccgccagtcttacaagaagggcatcatccggaagcca 1378
          | ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1335 tacgacaagctgagccgctccatccgccagtattacaagaagggcatcatccggaagcca 1394

Qy   1379 gacatctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccaggg 1438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1395 gacatctcccagcgcctcgtctaccagttcgtgcaccccatctgagtgcctggcccaggg 1454

Qy   1439 cctgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgaga 1498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1455 cctgaaacccgccctcaggggcctctctcctgcctgccctgcctcagccaggccctgaga 1514

Qy   1499 tgggggaaaacgggcagtgtgctctgctgctctgaccttccagagcccaaggtcagggag 1558
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1515 tgggggaaaacgggcagtctgctctgctgctctgaccttccagagcccaaggtcagggag 1574

Qy   1559 gggcaaccaactgccccaggggatatgggtcctctggggccttcgggaccatggggcag 1618
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   1575 gggcaaccaactgccccaggggatatgggtcctctggggccttcgggaccctggggcag 1634

Qy   1619 gggtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctc 1678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1635 gggtgcttcctcctcaggcccagctgctcccctggaggacagagggagacagggctgctc 1694

Qy   1679 cccaacacctgcctctgaccccagcatttccagagcagagcctacagaagggcagtgact 1738
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1695 cccaacacctgcctctgaccccagcatttccagagcagatcctacagaagggcagtgact 1754

Qy   1739 cgacaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctg 1798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1755 cgacaaaggccacaggcagtccaggcctctctctgctccatccccctgcctcccattctg 1814

Qy   1799 caccacacctggcatggtgcagggagacatctgcacccctcagttgggcagccaggagtg 1858
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1815 caccacacctggcatggtgcagggagacatctgcacccctgagttgggcagccaggagtg 1874

Qy   1859 cccccgggaatggataataaagatactag 1887
          |||||||||||||||||||||||||||||
Db   1875 cccccgggaatggataataaagatactag 1903


Search completed: August  6, 2002, 20:49:31
```

Job time: 9120 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 19:53:21 ; Search time 3498.6 Seconds
(without alignments)
11328.777 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 705160

Minimum DB seq length: 10
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	34	1.8	34	6	AX048158	AX048158 Sequence
2	31	1.6	31	6	AX048157	AX048157 Sequence
3	23	1.2	23	6	AX175520	AX175520 Sequence
4	23	1.2	23	6	AX175524	AX175524 Sequence
5	23	1.2	23	6	AX175525	AX175525 Sequence
6	23	1.2	23	6	AX175526	AX175526 Sequence
7	22.8	1.2	50	6	AX199688	AX199688 Sequence
8	22.6	1.2	48	6	AX221945	AX221945 Sequence
9	21.6	1.1	31	6	AX249190	AX249190 Sequence
c 10	21.4	1.1	49	10	MUSJHCD6	M64562 Mouse JH-Cd
11	21.2	1.1	45	6	A05116	A05116 Oligonucleo
12	21.2	1.1	48	9	HSTCK6X45	Z69575 H.sapiens m
13	21.2	1.1	48	9	HSTRK6X45	Z69547 H.sapiens m
c 14	21.2	1.1	50	6	AX250311	AX250311 Sequence
15	21	1.1	21	6	AX175518	AX175518 Sequence
c 16	21	1.1	37	6	AX139078	AX139078 Sequence
17	21	1.1	41	6	AX327072	AX327072 Sequence
18	20.8	1.1	48	6	AX221416	AX221416 Sequence
19	20.8	1.1	48	6	AX221427	AX221427 Sequence
20	20.8	1.1	48	6	AX274492	AX274492 Sequence
21	20.8	1.1	49	6	AX279693	AX279693 Sequence
c 22	20.8	1.1	50	6	AR123173	AR123173 Sequence
c 23	20.8	1.1	50	6	AX165853	AX165853 Sequence
24	20.6	1.1	48	6	AX221428	AX221428 Sequence
c 25	20.6	1.1	48	6	AX320747	AX320747 Sequence
26	20.6	1.1	50	9	HUMTCRVJ46	L39584 Homo sapien
27	20.4	1.1	48	6	AX221643	AX221643 Sequence
28	20.4	1.1	48	6	AX221644	AX221644 Sequence
29	20.4	1.1	48	6	AX221665	AX221665 Sequence
30	20.4	1.1	48	6	AX274490	AX274490 Sequence
31	20.4	1.1	48	6	AX274552	AX274552 Sequence
32	20.4	1.1	50	6	AR003562	AR003562 Sequence
33	20.4	1.1	50	6	AX057285	AX057285 Sequence
c 34	20.2	1.1	37	6	A46812	A46812 Sequence 9
35	20.2	1.1	42	6	AR021417	AR021417 Sequence
36	20.2	1.1	42	6	AR042979	AR042979 Sequence
37	20.2	1.1	42	6	AR161282	AR161282 Sequence
38	20.2	1.1	42	6	I43961	I43961 Sequence 50
39	20.2	1.1	42	6	I62970	I62970 Sequence 67
40	20.2	1.1	42	6	I88723	I88723 Sequence 67
41	20.2	1.1	45	9	S63845	S63845 (D3(TM4-del
c 42	20.2	1.1	46	6	AR140182	AR140182 Sequence
c 43	20.2	1.1	46	6	AR173249	AR173249 Sequence
c 44	20	1.1	36	6	AX007193	AX007193 Sequence
c 45	20	1.1	36	6	AX007365	AX007365 Sequence

ALIGNMENTS

RESULT 1
AX048158/c
LOCUS AX048158 34 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 11 from Patent WO0070092.
ACCESSION AX048158
VERSION AX048158.1 GI:11876974
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 34)
AUTHORS Kaufmann,J. and Xin,H.
TITLE Expression of ets-domain proteins in cancer
JOURNAL Patent: WO 0070092-A 11 23-NOV-2000;
Chiron Corporation (US)
FEATURES Location/Qualifiers
source 1. .34
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="GSEF-specific primer"
BASE COUNT 5 a 10 c 7 g 12 t
ORIGIN

```
  Query Match             1.8%;  Score 34;  DB 6;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 7.1e+04;
  Matches   34;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1850 ccaggagtgccccccgggaatggataataaagata 1883
          ||||||||||||||||||||||||||||||||||
Db     34 CCAGGAGTGCCCCCGGGAATGGATAATAAAGATA 1


RESULT   2
AX048157
LOCUS       AX048157                  31 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 10 from Patent WO0070092.
ACCESSION   AX048157
VERSION     AX048157.1  GI:11876973
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Kaufmann,J. and Xin,H.
  TITLE     Expression of ets-domain proteins in cancer
  JOURNAL   Patent: WO 0070092-A 10 23-NOV-2000;
            Chiron Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="GSEF-specific primer"
BASE COUNT        8 a      9 c     13 g      1 t
ORIGIN


  Query Match             1.6%;  Score 31;  DB 6;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1552 cagggaggggcaaccaactgccccaggggga 1582
          |||||||||||||||||||||||||||||||
Db      1 CAGGGAGGGGCAACCAACTGCCCCAGGGGGA 31


RESULT   3
AX175520
LOCUS       AX175520                  23 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 5 from Patent WO0142472.
ACCESSION   AX175520
VERSION     AX175520.1  GI:14598859
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Harkins,R., Lin,R.J., Luke,M., Monteclaro,F., Parkes,D., Parry,G.,
            Steinbrecher,R., van Heuit,P.T. and Xuan,J.A.
  TITLE     Dna encoding a novel prost-ets polypeptide
  JOURNAL   Patent: WO 0142472-A 5 14-JUN-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="probe"
BASE COUNT        4 a     12 c      2 g      5 t
ORIGIN


  Query Match             1.2%;  Score 23;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+06;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1783 cctgcctcccattctgcaccaca 1805
          |||||||||||||||||||||||
Db      1 CCTGCCTCCCATTCTGCACCACA 23


RESULT   4
AX175524
LOCUS       AX175524                  23 bp    mRNA    linear   PAT 03-JUL-2001
DEFINITION  Sequence 9 from Patent WO0142472.
ACCESSION   AX175524
VERSION     AX175524.1  GI:14598860
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Harkins,R., Lin,R.J., Luke,M., Monteclaro,F., Parkes,D., Parry,G.,
            Steinbrecher,R., van Heuit,P.T. and Xuan,J.A.
  TITLE     Dna encoding a novel prost-ets polypeptide
  JOURNAL   Patent: WO 0142472-A 9 14-JUN-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="antisense"
BASE COUNT        4 a      8 c      5 g      6 t
ORIGIN


  Query Match             1.2%;  Score 23;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+06;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1161 tccacctgtggcagttcctcaag 1183
          |||||||||||||||||||||||
Db      1 TCCACCTGTGGCAGTTCCTCAAG 23


RESULT   5
AX175525
LOCUS       AX175525                  23 bp    mRNA    linear   PAT 03-JUL-2001
DEFINITION  Sequence 10 from Patent WO0142472.
ACCESSION   AX175525
VERSION     AX175525.1  GI:14598861
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Harkins,R., Lin,R.J., Luke,M., Monteclaro,F., Parkes,D., Parry,G.,
            Steinbrecher,R., van Heuit,P.T. and Xuan,J.A.
  TITLE     Dna encoding a novel prost-ets polypeptide
  JOURNAL   Patent: WO 0142472-A 10 14-JUN-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="antisense"
BASE COUNT        8 a      7 c      6 g      2 t
ORIGIN


  Query Match             1.2%;  Score 23;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+06;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1734 tgactcgacaaaggccacaggca 1756
          |||||||||||||||||||||||
Db      1 TGACTCGACAAAGGCCACAGGCA 23


RESULT   6
```

```
AX175526
LOCUS       AX175526                 23 bp    mRNA    linear   PAT 03-JUL-2001
DEFINITION  Sequence 11 from Patent WO0142472.
ACCESSION   AX175526
VERSION     AX175526.1  GI:14598862
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Harkins,R., Lin,R.J., Luke,M., Monteclaro,F., Parkes,D., Parry,G.,
            Steinbrecher,R., van Heuit,P.T. and Xuan,J.A.
  TITLE     Dna encoding a novel prost-ets polypeptide
  JOURNAL   Patent: WO 0142472-A 11 14-JUN-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="antisense"
BASE COUNT        6 a      8 c      5 g      4 t
ORIGIN


  Query Match             1.2%;  Score 23;  DB 6;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+06;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1699 ccagcatttccagagcagagcct 1721
          |||||||||||||||||||||||
Db      1 CCAGCATTTCCAGAGCAGAGCCT 23


RESULT   7
AX199688
LOCUS       AX199688                 50 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 618 from Patent WO0151670.
ACCESSION   AX199688
VERSION     AX199688.1  GI:15390129
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.D.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0151670-A 618 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg44009326"
     misc_feature    26
                     /note="2 of 2 allelic variants (617 is other entry)"
BASE COUNT        4 a     12 c     25 g      9 t
ORIGIN


  Query Match             1.2%;  Score 22.8;  DB 6;  Length 50;
  Best Local Similarity   66.0%;  Pred. No. 4.8e+06;
  Matches   33;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    485 tcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagag-534
          |||||| |||||| || |     ||| ||||| |  || |||  ||  ||
Db      1 TCGCGGCCAGGCTGGGCGCGCAGGGCGGCGGGTGGTGTTGGTTACGTCAG 50


RESULT   8
```

```
AX221945
LOCUS       AX221945                 48 bp    mRNA    linear   PAT 07-SEP-2001
DEFINITION  Sequence 7387 from Patent WO0159103.
ACCESSION   AX221945
VERSION     AX221945.1  GI:15549669
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Blatt,L., Mcswiggen,J. and Chowrira,B.M.
  TITLE     Method and reagent for the modulation and diagnosis of cd20 and
            nogo gene expression
  JOURNAL   Patent: WO 0159103-A 7387 16-AUG-2001;
            RIBOZYME PHARMACEUTICALS, INC. (US) ; Blatt, Lawrence (US) ;
            McSwiggen, James (US) ; Chowrira, Bharat M. (US)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Nucleic Acid"
BASE COUNT       10 a     13 c     18 g      7 t
ORIGIN


  Query Match             1.2%;  Score 22.6;  DB 6;  Length 48;
  Best Local Similarity   68.9%;  Pred. No. 5.3e+06;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    137 ggctgggggtaggggactccctacaggcacgcagccctgagaccg 181
          ||||||| | |||  ||||||| || | ||   | || | ||| |
Db      1 GGCTGGGCGGAGGAAACTCCCTTCAAGGACATCGTCCGGGGACTG 45


RESULT   9
AX249190
LOCUS       AX249190                 31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 1269 from Patent WO0166800.
ACCESSION   AX249190
VERSION     AX249190.1  GI:15863813
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 1269 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        5 a     11 c      6 g      8 t      1 others
ORIGIN


  Query Match             1.1%;  Score 21.6;  DB 6;  Length 31;
  Best Local Similarity   80.0%;  Pred. No. 8.6e+06;
  Matches   24;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy     59 ccagtcctccaagcctgctgccagctccct 88
          ||||| ||||||||||:| || || || |||
Db      1 CCAGTGCTCCAAGCCYGGTGTCATCTTCCT 30


RESULT  10
MUSJHCD6/c
LOCUS       MUSJHCD6                 49 bp    DNA     linear   ROD 27-APR-1993
DEFINITION  Mouse JH-Cdelta locus sequence.
ACCESSION   M64562
VERSION     M64562.1  GI:198478
```

```
KEYWORDS    C-region; immunoglobulin; immunoglobulin heavy chain; switch
            recombination junction.
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) DNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Owens,J.D.Jr., Finkelman,F.D., Mountz,J.D. and Mushinski,J.F.
  TITLE     Nonhomologous recombination at sites within the mouse JH-Cdelta
            locus accompanies Cmu deletion and switch to immunoglobulin D
            secretion
  JOURNAL   Mol. Cell. Biol. 11, 5660-5670 (1991)
  MEDLINE   92017847
COMMENT     This reference reports bases 1 to 25.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /cell_line="KWD8"
                     /cell_type="Hybridoma"
BASE COUNT        7 a     17 c     14 g     11 t
ORIGIN


  Query Match             1.1%;  Score 21.4;  DB 10;  Length 49;
  Best Local Similarity   66.0%;  Pred. No. 8.3e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    277 cccagatgcctggctgagacacgccagtggcctcagctgcccacacc 323
          || |||||   ||||| ||||   ||| |    | |  |||||||||
Db     49 CCAAGATGGTGGGCTGGGACAGAGCAGAGCTTGCTGTGGCCCACACC 3


RESULT  11
A05116
LOCUS       A05116                   45 bp    DNA     linear   PAT 04-MAY-1993
DEFINITION  Oligonucleotide U7 for bovine growth hormone.
ACCESSION   A05116
VERSION     A05116.1  GI:345003
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   .
  TITLE     METHOD FOR THE PRODUCTION OF BOVINE GROWTH HORMONE USING A
            SYNTHETIC GENE
  JOURNAL   Patent: WO 8805078-A 13 14-JUL-1988;
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        8 a     17 c     10 g     10 t
ORIGIN


  Query Match             1.1%;  Score 21.2;  DB 6;  Length 45;
  Best Local Similarity   76.5%;  Pred. No. 9.1e+06;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1691 ctctgaccccagcatttccagagcagagcctaca 1724
           ||| | |||| ||| ||||  ||||||| || ||
Db      1 CTCGGGCCCCTGCAGTTCCTCAGCAGAGTCTTCA 34


RESULT  12
HSTCK6X45
LOCUS       HSTCK6X45                48 bp    mRNA    linear   PRI 05-FEB-1997
DEFINITION  H.sapiens mRNA for T cell receptor beta chain junctional region
            (clone K6-45).
ACCESSION   Z69575
VERSION     Z69575.1  GI:1770552
KEYWORDS    diversity region; joining region; T cell receptor; variable region.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Naserke,H.E., Durinovic-Bello,I., Seidel,D. and Ziegler,A.G.
  TITLE     The T-cell receptor beta chain CDR3 region of BV8S1/BJ1S5
            transcripts in type 1 diabetes
  JOURNAL   Immunogenetics 45 (2), 87-96 (1996)
  MEDLINE   97113296
REFERENCE   2  (bases 1 to 48)
  AUTHORS   Naserke,H.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1996) Naserke H.E., Diabetes Research Institute,
            Koelner Platz 1, D-80804 Muenchen, FRG
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /isolate="healthy control WS"
                     /db_xref="taxon:9606"
                     /tissue_type="peripheral blood"
                     /cell_type="lymphocytes"
                     /clone="K6-45"
                     /note="HLA type: HLA-A31, 1; Cw3, 3; B62, 55; DRB1*1301,
                     0401; DRB3*0202, -; DQA1*0103, 0301; DQB1*0603, 0302;
                     DPA1*-, 01; DPB1*0301, 0201"
     gene            1. .11
                     /gene="V beta 8.1"
     V_segment       <1. .11
                     /gene="V beta 8.1"
     D_segment       12. .28
                     /note="N-D-N region"
     J_segment       29. .>48
                     /gene="J beta 1.5"
     gene            29. .48
                     /gene="J beta 1.5"
BASE COUNT        8 a     13 c     16 g     11 t
ORIGIN


  Query Match             1.1%;  Score 21.2;  DB 9;  Length 48;
  Best Local Similarity   69.0%;  Pred. No. 9e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    651 gggccagcagtcgggaggagccacctgaccagcctgagcagt 692
          | |||||||| |||| || |  |   || |||||  |||| |
Db      2 GTGCCAGCAGCCGGGCGGCGTTAATGGATCAGCCCCAGCATT 43


RESULT  13
HSTRK6X45
LOCUS       HSTRK6X45                48 bp    mRNA    linear   PRI 05-FEB-1997
DEFINITION  H.sapiens mRNA for T cell receptor beta chain junctional region
            (clone K6-45).
ACCESSION   Z69547
VERSION     Z69547.1  GI:1770714
KEYWORDS    diversity region; joining region; T cell receptor; variable region.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Naserke,H.E., Durinovic-Bello,I., Seidel,D. and Ziegler,A.G.
  TITLE     The T-cell receptor beta chain CDR3 region of BV8S1/BJ1S5
            transcripts in type 1 diabetes
  JOURNAL   Immunogenetics 45 (2), 87-96 (1996)
  MEDLINE   97113296
REFERENCE   2  (bases 1 to 48)
  AUTHORS   Naserke,H.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1996) Naserke H.E., Diabetes Research Institute,
```

```
          Koelner Platz 1, D-80804 Muenchen, FRG
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /isolate="healthy control WS"
                     /db_xref="taxon:9606"
                     /tissue_type="peripheral blood"
                     /cell_type="lymphocytes"
                     /clone="K6-45"
                     /note="HLA type: HLA-A31, 1; Cw3, 3; B62, 55; DRB1*1301,
                     0401; DRB3*0202, -; DQA1*0103, 0301; DQB1*0603, 0302;
                     DPA1*-, 01; DPB1*0301, 0201"
     gene            1. .11
                     /gene="V beta 8.1"
     V_segment       <1. .11
                     /gene="V beta 8.1"
     D_segment       12. .28
                     /note="N-D-N region"
     J_segment       29. .>48
                     /gene="J beta 1.5"
     gene            29. .48
                     /gene="J beta 1.5"
BASE COUNT        8 a     13 c     16 g     11 t
ORIGIN


  Query Match             1.1%;  Score 21.2;  DB 9;  Length 48;
  Best Local Similarity   69.0%;  Pred. No. 9e+06;
  Matches   29;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    651 gggccagcagtcgggaggagccacctgaccagcctgagcagt 692
          | |||||||| |||| || |  |   || |||||  |||| |
Db      2 GTGCCAGCAGCCGGGCGGCGTTAATGGATCAGCCCCAGCATT 43


RESULT  14
AX250311/c
LOCUS       AX250311                 50 bp    DNA     linear   PAT 05-OCT-2001
DEFINITION  Sequence 51 from Patent WO0168867.
ACCESSION   AX250311
VERSION     AX250311.1  GI:15984079
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Leadlay,P.F., Staunton,J. and Oliynyk,M.C.
  TITLE     Polyketides and their synthesis
  JOURNAL   Patent: WO 0168867-A 51 20-SEP-2001;
            Biotica Technology Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT        5 a     20 c     17 g      8 t
ORIGIN


  Query Match             1.1%;  Score 21.2;  DB 6;  Length 50;
  Best Local Similarity   76.5%;  Pred. No. 8.9e+06;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     27 gcactctgccgtgtccacactgccccacagaccc 60
          | |||  |||||||| |||||| || || | |||
Db     43 GGACTGCGCCGTGTCGACACTGACCGACGGGCCC 10


RESULT  15
AX175518
LOCUS       AX175518                 21 bp    DNA     linear   PAT 03-JUL-2001
DEFINITION  Sequence 3 from Patent WO0142472.
ACCESSION   AX175518
VERSION     AX175518.1  GI:14598857
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Harkins,R., Lin,R.J., Luke,M., Monteclaro,F., Parkes,D., Parry,G.,
            Steinbrecher,R., van Heuit,P.T. and Xuan,J.A.
  TITLE     Dna encoding a novel prost-ets polypeptide
  JOURNAL   Patent: WO 0142472-A 3 14-JUN-2001;
            SCHERING AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer"
BASE COUNT        8 a      6 c      5 g      2 t
ORIGIN


  Query Match             1.1%;  Score 21;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+07;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1732 agtgactcgacaaaggccaca 1752
          |||||||||||||||||||||
Db      1 AGTGACTCGACAAAGGCCACA 21


Search completed: August  6, 2002, 22:33:25
Job time: 9604 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 20:55:46 ; Search time 311.92 Seconds
(without alignments)
10425.224 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 1900032

Minimum DB seq length: 10
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS5/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	34	1.8	34	22	AAC83265	GSEF specific PCR
	2	31	1.6	31	22	AAC83264	GSEF specific PCR
c	3	24.6	1.3	42	20	AAX35839	Antisense oligonuc
	4	23.6	1.2	47	21	AAZ65828	Human map-related
c	5	23.6	1.2	50	22	AAL34612	Human SNP oligonuc
	6	23	1.2	23	22	AAD07340	Human transcriptio
	7	23	1.2	23	22	AAD07341	Antisense oligo #1
	8	23	1.2	23	22	AAD07342	Antisense oligo #2
	9	23	1.2	23	22	AAD07343	Antisense oligo #3
c	10	22.8	1.2	50	22	AAL31921	Human SNP oligonuc
	11	22.8	1.2	50	22	AAH89837	Human coding sequ
	12	22.6	1.2	48	23	ABK07387	Human NOGO Amberzy
	13	22.4	1.2	50	22	AAL28314	Human SNP oligonuc
c	14	22.4	1.2	50	22	AAL34308	Human SNP oligonuc
	15	22	1.2	45	22	AAD19606	Mutagenic PCR prim
c	16	22	1.2	45	22	AAD19607	Mutagenic PCR prim
	17	22	1.2	45	22	AAF74532	Human MKK3 site-di
c	18	22	1.2	45	22	AAF74533	Human MKK3 site-di
	19	22	1.2	45	22	AAC64157	Human MKK mutageni
c	20	22	1.2	45	22	AAC64158	Human MKK mutageni
	21	21.8	1.2	49	18	AAT80517	Hepatoma AS-30D Ty
	22	21.8	1.2	50	22	AAL32325	Human SNP oligonuc
	23	21.6	1.1	45	22	AAF60322	PDGF-B receptor bi
	24	21.4	1.1	39	12	AAQ11438	Probe #4 complemen
	25	21.2	1.1	36	20	AAZ23808	Human Kv6.2 DNA fr
	26	21.2	1.1	50	22	AAL29118	Human SNP oligonuc
	27	21	1.1	21	22	AAD07338	Human transcriptio
c	28	21	1.1	37	22	AAF56988	FIV gene cloning r
	29	21	1.1	50	22	AAL34241	Human SNP oligonuc
	30	20.8	1.1	29	21	AAA04697	Polymorphic fragme
	31	20.8	1.1	48	23	ABK06858	Human NOGO Amberzy
	32	20.8	1.1	48	23	ABK06869	Human NOGO Amberzy
	33	20.8	1.1	49	23	ABA10736	Tail adaptor oligo
c	34	20.8	1.1	50	17	AAT40111	Human Mer receptor
	35	20.8	1.1	50	19	AAV39429	Humanised anti-HM1
c	36	20.8	1.1	50	19	AAV39430	Humanised anti-HM1
	37	20.8	1.1	50	20	AAX59472	Primer used in con
c	38	20.8	1.1	50	20	AAX59473	Primer used in con
c	39	20.8	1.1	50	23	ABL01057	Human SNP involvin
	40	20.6	1.1	48	23	ABK06870	Human NOGO Amberzy
	41	20.4	1.1	31	22	AAI30781	Human single nucle
	42	20.4	1.1	45	21	AAC81266	Human/mouse histon
c	43	20.4	1.1	47	21	AAZ65894	Human map-related
	44	20.4	1.1	48	23	ABK07085	Human NOGO Amberzy
	45	20.4	1.1	48	23	ABK07086	Human NOGO Amberzy

ALIGNMENTS

RESULT 1
AAC83265/c
ID AAC83265 standard; DNA; 34 BP.
XX
AC AAC83265;
XX
DT 16-MAR-2001 (first entry)
XX
DE GSEF specific PCR primer SEQ ID 11.
XX
KW Transcription factor; gland-specific Ets transcription factor; GSEF;
KW metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200070092-A1.
XX
PD 23-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US13173.
XX
PR 14-MAY-1999; 99US-0134112.
XX
PA (CHIR) CHIRON CORP.
XX
PI Kaufmann J, Xin H, Harrowe G;
XX
DR WPI; 2001-041019/05.
XX
PT Detecting metastatic and potential metastatic cancerous cells, useful

PT for diagnosing, prognosing, grading and staging of cancers by detecting
PT gland-specific Ets transcription factor gene product in a biological
PT sample from a cell -
XX
PS Example 7; Page 56; 95pp; English.
XX
CC This invention relates to a method for the detection and determination of
CC the metastatic potential of a cell. The method comprises detecting a
CC gland-specific Ets transcription factor (GSEF) gene product in a test
CC sample. Detection of a GSEF gene product in the test sample in amount
CC lower than that in a normal cell, is indicative of a cell with high
CC metastatic potential. The method is useful for determining the metastatic
CC potential of a cell, for the diagnosis and prognosis of cancer as well as
CC grading and staging of cancers by detecting GSEF expression in a
CC biological test sample. The method may also be used to monitor patients
CC having a predisposition to develop a particular cancer. GSEF polypeptides
CC are useful for producing antibodies, in cancer diagnosis, prognosis,
CC grading, staging and management of breast and prostate tumours, and in
CC detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC useful in gene therapy. GSEF gene product expression levels can be used
CC in conjunction with any tissue in which an alteration in GSEF gene
CC product expression levels is associated with development of a
CC cancer-associated phenotype. Cancers, which can be monitored include
CC cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC represents a PCR primer specific for the human GSEF DNA sequence. The
CC primer is used for the chromosamal mapping of the GSEF gene.
XX
SQ Sequence 34 BP; 5 A; 10 C; 7 G; 12 T; 0 other;

Query Match 1.8%; Score 34; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1850 ccaggagtgccccccgggaatggataataaagata 1883
||||||||||||||||||||||||||||||||||
Db 34 CCAGGAGTGCCCCCGGGAATGGATAATAAAGATA 1

RESULT 2
AAC83264
ID AAC83264 standard; DNA; 31 BP.
XX
AC AAC83264;
XX
DT 16-MAR-2001 (first entry)
XX
DE GSEF specific PCR primer SEQ ID 10.
XX
KW Transcription factor; gland-specific Ets transcription factor; GSEF;
KW metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200070092-A1.
XX
PD 23-NOV-2000.
XX
PF 12-MAY-2000; 2000WO-US13173.
XX
PR 14-MAY-1999; 99US-0134112.
XX
PA (CHIR) CHIRON CORP.
XX
PI Kaufmann J, Xin H, Harrowe G;
XX
DR WPI; 2001-041019/05.
XX

PT Detecting metastatic and potential metastatic cancerous cells, useful
PT for diagnosing, prognosing, grading and staging of cancers by detecting
PT gland-specific Ets transcription factor gene product in a biological
PT sample from a cell -
XX
PS Example 7; Page 56; 95pp; English.
XX
CC This invention relates to a method for the detection and determination of
CC the metastatic potential of a cell. The method comprises detecting a
CC gland-specific Ets transcription factor (GSEF) gene product in a test
CC sample. Detection of a GSEF gene product in the test sample in amount
CC lower than that in a normal cell, is indicative of a cell with high
CC metastatic potential. The method is useful for determining the metastatic
CC potential of a cell, for the diagnosis and prognosis of cancer as well as
CC grading and staging of cancers by detecting GSEF expression in a
CC biological test sample. The method may also be used to monitor patients
CC having a predisposition to develop a particular cancer. GSEF polypeptides
CC are useful for producing antibodies, in cancer diagnosis, prognosis,
CC grading, staging and management of breast and prostate tumours, and in
CC detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC useful in gene therapy. GSEF gene product expression levels can be used
CC in conjunction with any tissue in which an alteration in GSEF gene
CC product expression levels is associated with development of a
CC cancer-associated phenotype. Cancers, which can be monitored include
CC cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC represents a PCR primer specific for the human GSEF DNA sequence. The
CC primer is used for the chromosamal mapping of the GSEF gene.
XX
SQ Sequence 31 BP; 8 A; 9 C; 13 G; 1 T; 0 other;

Query Match 1.6%; Score 31; DB 22; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1552 cagggaggggcaaccaactgccccaggggga 1582
|||||||||||||||||||||||||||||||
Db 1 cagggaggggcaaccaactgccccaggggga 31

RESULT 3
AAX35839/c
ID AAX35839 standard; DNA; 42 BP.
XX
AC AAX35839;
XX
DT 14-JUL-1999 (first entry)
XX
DE Antisense oligonucleotide for mammalian ACE mRNA.
XX
KW Antisense oligonucleotide; mammalian; angiotensin converting enzyme;
KW ACE; hypertension; ss.
XX
OS Synthetic.
XX
PN WO9915643-A2.
XX
PD 01-APR-1999.
XX
PF 25-SEP-1998; 98WO-US20121.
XX
PR 25-SEP-1997; 97US-0059661.
XX
PA (UYFL) UNIV FLORIDA.
XX
PI Mohuczy D, Phillips MI;
XX
DR WPI; 1999-276982/23.
XX
PT New antisense oligonucleotide compositions

XX
PS Claim 4; Page 16; 145pp; English.
XX
CC AAX35833-49 represent antisense oligonucleotides for mammalian
CC angiotensin converting enzyme (ACE) mRNA. The antisense
CC oligonucleotides can be used for treating or preventing
CC hypertension. The antisense oligonucleotides can also be
CC labeled and used for visualising ACE mRNA in a cell, and for
CC producing transgenic animals.
XX
SQ Sequence 42 BP; 1 A; 17 C; 14 G; 10 T; 0 other;

Query Match 1.3%; Score 24.6; DB 20; Length 42;
Best Local Similarity 76.9%; Pred. No. 2.9e+04;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 392 gccgccagcccaaacagcagcggcatgggcagcgccagc 430
| || |||| |||||||||||||| |||||| |||| |
Db 39 GGCGGCAGCAGCAACAGCAGCGGCAGCGGCAGCAGCAGC 1

RESULT 4
AAZ65828
ID AAZ65828 standard; DNA; 47 BP.
XX
AC AAZ65828;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human map-related biallelic marker SEQ ID NO:175.
XX
KW Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation;
KW diagnosis; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(24,G)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO9954500-A2.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1999; 99WO-IB00822.
XX
PR 21-APR-1998; 98US-0082614.
PR 23-NOV-1998; 98US-0109732.
XX
PA (GEST) GENSET.
XX
PI Cohen D, Blumenfeld M, Chumakov I;
XX
DR WPI; 2000-013267/01.
XX
PT Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome -
XX
PS Claim 1; Page 267; 2745pp; English.
XX
CC AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the
CC invention have a variety of uses: they can be used for high density
CC mapping of the human genome, and in complex association studies and
CC haplotyping studies which are useful in determining the genetic basis
CC for disease states. Compositions and methods of the invention can also
CC be useful for the identification of the targets for the development of
CC pharmaceutical agents and diagnostic methods, as well as the
CC characterisation of the differential efficacious responses to and side
CC effects from pharmaceutical agents acting on a disease as well as other
CC treatment.
CC N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297
CC and 3367, are not actually given a sequence in the Sequence Listing
CC from the present invention.
XX
SQ Sequence 47 BP; 6 A; 14 C; 16 G; 11 T; 0 other;

Query Match 1.2%; Score 23.6; DB 21; Length 47;
Best Local Similarity 69.6%; Pred. No. 4.8e+04;
Matches 32; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1484 agccaggccctgagatgggggaaaacgggcagtgtgctctgctgct 1529
|| || |||||| || || | ||| |||||||||||| |
Db 1 aggcaagccctgtaccctcgggaagccggctgtgtgctctgctgtt 46

RESULT 5
AAL34612/c
ID AAL34612 standard; DNA; 50 BP.
XX
AC AAL34612;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #7820.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3652; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune

```
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 8 A; 17 C; 18 G; 7 T; 0 other;


  Query Match             1.2%;  Score 23.6;  DB 22;  Length 50;
  Best Local Similarity   69.6%;  Pred. No. 4.8e+04;
  Matches   32;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    273 gccccccagatgcctggctgagacacgccagtggcctcagctgccc 318
          |||||||||  | ||||| | | ||||  |||| || | |  | ||
Db     47 GCCCCCCAGGCGGCTGGCAGTGCCACGGAAGTGTCCACTGAGGTCC 2


RESULT   6
AAD07340
ID   AAD07340 standard; DNA; 23 BP.
XX
AC   AAD07340;
XX
DT   04-AUG-2001  (first entry)
XX
DE   Human transcription factor, prost-ets mRNA expressing Taqman probe.
XX
KW   Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;
KW   vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;
KW   antisense therapy; ribozyme therapy; therapeutic; cell proliferation;
KW   probe; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag= a
FT                   /mod_base= OTHER
FT                   /note= "FAM-labelled Cytosine"
FT   modified_base   23
FT                   /*tag= a
FT                   /mod_base= OTHER
FT                   /note= "TAMRA-labelled Adenosine"
XX
PN   WO200142472-A2.
XX
PD   14-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-US42295.
XX
PR   30-NOV-1999;   99US-0168182.
PR   21-NOV-2000; 2000US-0718159.
XX
PA   (SCHD ) SCHERING AG.
XX
PI   Harkins R,  Lin RJ,  Luke M,  Monteclaro F,  Parkes D,  Parry G;
PI   Steinbrecher R,  Van Heuit PT,  Xuan J;
XX
DR   WPI; 2001-381692/40.
XX
PT   New transcription factor polypeptides, designated PROST-Ets
PT   polypeptides, useful for diagnostic, research and therapeutic
PT   applications -
XX
PS   Example 2; Page 48; 74pp; English.
XX
CC   The invention relates to human transcription factor protein, designated
CC   as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and
CC   nucleic acids are useful for treating a disease-state in human patient,
CC   such as prostate, breast and ovarian cancer which is associated with
CC   inappropriate expression of PROST-Ets and where the patient is in need of
CC   decreased levels or activity of PROST-Ets. PROST-Ets is useful for
CC   diagnosing and treating diseases of cell proliferation, such as prostate
CC   cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used
CC   as DNA probes, as target for antisense and ribozyme therapy, as template
CC   for producing antisense polynucleotides, and as diagnostic reagent for
CC   research, biological, clinical and therapeutic purposes. Prost-ets DNA is
CC   also used in gene therapy. PROST-Ets protein is useful for generating
CC   antibodies, and for stimulating immune response in cells containing
CC   PROST-Ets proteins. The present DNA sequence is Taqman probe which is
CC   used for the expression of human transcription factor, prost-ets mRNA.
XX
SQ   Sequence 23 BP; 4 A; 12 C; 2 G; 5 T; 0 other;


  Query Match             1.2%;  Score 23;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 5.6e+04;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1783 cctgcctcccattctgcaccaca 1805
          |||||||||||||||||||||||
Db      1 cctgcctcccattctgcaccaca 23


RESULT   7
AAD07341
ID   AAD07341 standard; RNA; 23 BP.
XX
AC   AAD07341;
XX
DT   04-AUG-2001  (first entry)
XX
DE   Antisense oligo #1 to evaluate human prost-ets in cell proliferation.
XX
KW   Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;
KW   vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;
KW   antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142472-A2.
XX
PD   14-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-US42295.
XX
PR   30-NOV-1999;   99US-0168182.
PR   21-NOV-2000; 2000US-0718159.
XX
PA   (SCHD ) SCHERING AG.
XX
PI   Harkins R,  Lin RJ,  Luke M,  Monteclaro F,  Parkes D,  Parry G;
PI   Steinbrecher R,  Van Heuit PT,  Xuan J;
XX
DR   WPI; 2001-381692/40.
XX
PT   New transcription factor polypeptides, designated PROST-Ets
PT   polypeptides, useful for diagnostic, research and therapeutic
PT   applications -
XX
PS   Claim 36; Page 56; 74pp; English.
XX
CC   The invention relates to human transcription factor protein, designated
CC   as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and
CC   nucleic acids are useful for treating a disease-state in human patient,
CC   such as prostate, breast and ovarian cancer which is associated with
CC   inappropriate expression of PROST-Ets and where the patient is in need of
CC   decreased levels or activity of PROST-Ets. PROST-Ets is useful for
CC   diagnosing and treating diseases of cell proliferation, such as prostate
CC   cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used
CC   as DNA probes, as target for antisense and ribozyme therapy, as template
CC   for producing antisense polynucleotides, and as diagnostic reagent for
CC   research, biological, clinical and therapeutic purposes. Prost-ets DNA is
CC   also used in gene therapy. PROST-Ets protein is useful for generating
CC   antibodies, and for stimulating immune response in cells containing
```

CC PROST-Ets proteins. The present DNA sequence is antisense oligonucleotide
CC which is used for evaluating the role of human transcription factor,
CC PROST-Ets in cell proliferation to reduce Prost-ets mRNA expression in
CC the selected prostate tumour cell lines.
XX
SQ Sequence 23 BP; 4 A; 8 C; 5 G; 6 U; 0 other;

Query Match 1.2%; Score 23; DB 22; Length 23;
Best Local Similarity 73.9%; Pred. No. 5.6e+04;
Matches 17; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1161 tccacctgtggcagttcctcaag 1183
:|||||:|:||||||::||:||||
Db 1 uccaccuguggcaguuccucaag 23

RESULT 8
AAD07342
ID AAD07342 standard; RNA; 23 BP.
XX
AC AAD07342;
XX
DT 04-AUG-2001 (first entry)
XX
DE Antisense oligo #2 to evaluate human prost-ets in cell proliferation.
XX
KW Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;
KW vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;
KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.
XX
OS Homo sapiens.
XX
PN WO200142472-A2.
XX
PD 14-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US42295.
XX
PR 30-NOV-1999; 99US-0168182.
PR 21-NOV-2000; 2000US-0718159.
XX
PA (SCHD) SCHERING AG.
XX
PI Harkins R, Lin RJ, Luke M, Monteclaro F, Parkes D, Parry G;
PI Steinbrecher R, Van Heuit PT, Xuan J;
XX
DR WPI; 2001-381692/40.
XX
PT New transcription factor polypeptides, designated PROST-Ets
PT polypeptides, useful for diagnostic, research and therapeutic
PT applications -
XX
PS Claim 37; Page 56; 74pp; English.
XX
CC The invention relates to human transcription factor protein, designated
CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and
CC nucleic acids are useful for treating a disease-state in human patient,
CC such as prostate, breast and ovarian cancer which is associated with
CC inappropriate expression of PROST-Ets and where the patient is in need of
CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for
CC diagnosing and treating diseases of cell proliferation, such as prostate
CC cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used
CC as DNA probes, as target for antisense and ribozyme therapy, as template
CC for producing antisense polynucleotides, and as diagnostic reagent for
CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is
CC also used in gene therapy. PROST-Ets protein is useful for generating
CC antibodies, and for stimulating immune response in cells containing
CC PROST-Ets proteins. The present DNA sequence is antisense oligonucleotide
CC which is used for evaluating the role of human transcription factor,
CC PROST-Ets in cell proliferation to reduce Prost-ets mRNA expression in
CC the selected prostate tumour cell lines.
XX
SQ Sequence 23 BP; 8 A; 7 C; 6 G; 2 U; 0 other;

Query Match 1.2%; Score 23; DB 22; Length 23;
Best Local Similarity 91.3%; Pred. No. 5.6e+04;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1734 tgactcgacaaaggccacaggca 1756
:|||:||||||||||||||||||
Db 1 ugacucgacaaaggccacaggca 23

RESULT 9
AAD07343
ID AAD07343 standard; RNA; 23 BP.
XX
AC AAD07343;
XX
DT 04-AUG-2001 (first entry)
XX
DE Antisense oligo #3 to evaluate human prost-ets in cell proliferation.
XX
KW Human; transcription factor protein; PROST-Ets; cytostatic; gene therapy;
KW vaccine; cancer; prostate; breast; ovary; benign prostatic hyperplasia;
KW antisense therapy; ribozyme therapy; therapeutic; cell proliferation; ss.
XX
OS Homo sapiens.
XX
PN WO200142472-A2.
XX
PD 14-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US42295.
XX
PR 30-NOV-1999; 99US-0168182.
PR 21-NOV-2000; 2000US-0718159.
XX
PA (SCHD) SCHERING AG.
XX
PI Harkins R, Lin RJ, Luke M, Monteclaro F, Parkes D, Parry G;
PI Steinbrecher R, Van Heuit PT, Xuan J;
XX
DR WPI; 2001-381692/40.
XX
PT New transcription factor polypeptides, designated PROST-Ets
PT polypeptides, useful for diagnostic, research and therapeutic
PT applications -
XX
PS Claim 38; Page 56; 74pp; English.
XX
CC The invention relates to human transcription factor protein, designated
CC as PROST-Ets and its corresponding cDNA molecule. PROST-Ets proteins and
CC nucleic acids are useful for treating a disease-state in human patient,
CC such as prostate, breast and ovarian cancer which is associated with
CC inappropriate expression of PROST-Ets and where the patient is in need of
CC decreased levels or activity of PROST-Ets. PROST-Ets is useful for
CC diagnosing and treating diseases of cell proliferation, such as prostate
CC cancer and benign prostatic hyperplasia. Prost-ets nucleic acids are used
CC as DNA probes, as target for antisense and ribozyme therapy, as template
CC for producing antisense polynucleotides, and as diagnostic reagent for
CC research, biological, clinical and therapeutic purposes. Prost-ets DNA is
CC also used in gene therapy. PROST-Ets protein is useful for generating
CC antibodies, and for stimulating immune response in cells containing
CC PROST-Ets proteins. The present DNA sequence is antisense oligonucleotide
CC which is used for evaluating the role of human transcription factor,
CC PROST-Ets in cell proliferation to reduce Prost-ets mRNA expression in
CC the selected prostate tumour cell lines.
XX
SQ Sequence 23 BP; 6 A; 8 C; 5 G; 4 U; 0 other;

Query Match 1.2%; Score 23; DB 22; Length 23;
Best Local Similarity 82.6%; Pred. No. 5.6e+04;

```
  Matches   19;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1699 ccagcatttccagagcagagcct 1721
          ||||||:::|||||||||||||:
Db      1 ccagcauuuccagagcagagccu 23


RESULT  10
AAL31921/c
ID   AAL31921 standard; DNA; 50 BP.
XX
AC   AAL31921;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #5129.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000;  2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2862; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 8 A; 19 C; 15 G; 8 T; 0 other;


  Query Match             1.2%;  Score 22.8;  DB 22;  Length 50;
  Best Local Similarity   79.4%;  Pred. No. 7.1e+04;
  Matches   27;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   1189 gctactcaagccccacagctatggccgcttcatt 1222
           ||| || |||||||||||||| |||| |||  | |
Db     49 GCTGCTGAAGCCCCACAGCTGTGGCTGCTGAAGT 16


RESULT  11
AAH89837
ID   AAH89837 standard; DNA; 50 BP.
XX
AC   AAH89837;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human  coding sequence polymorphic site SEQ ID NO: 618.
XX
KW   Human; single nucleotide polymorphism; SNP; paternity test;
KW   forensic test; aberrant protein expression; ds.
XX
OS   Homo sapiens.
XX
PN   WO200151670-A2.
XX
PD   19-JUL-2001.
XX
PF   05-JAN-2001; 2001WO-US00322.
XX
PR   07-JAN-2000; 2000US-0174962.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach MD;
XX
DR   WPI; 2001-451871/48.
DR   P-PSDB; AAM00718.
XX
PT   Isolated human polynucleotides containing single nucleotide
PT   polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
PT   infection and diabetes -
XX
PS   Claim 1; Page 282; 475pp; English.
XX
CC   The present invention relates to human nucleic acids containing single
CC   nucleotide polymorphisms (SNPs). These can be used in forensic and
CC   paternity tests, and to aid in the treatment of diseases associated with
CC   aberrant protein expression, including cancer, amyloidosis, diabetes,
CC   Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
CC   glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,
CC   meningitis, muscular disorders, dementia, neurological diseases, tuberous
CC   sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
CC   osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
CC   autoimmunity. The present sequence is a polymorphism-containing
CC   oligonucleotide fragment of the invention.
XX
SQ   Sequence 50 BP; 4 A; 12 C; 25 G; 9 T; 0 other;


  Query Match             1.2%;  Score 22.8;  DB 22;  Length 50;
  Best Local Similarity   66.0%;  Pred. No. 7.1e+04;
  Matches   33;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    485 tcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagag 534
          |||||| |||||| || |     ||| ||||| |  || |||  ||  ||
Db      1 tcgcggccaggctgggcgcgcagggcggcgggtggtgttggttacgtcag 50


RESULT  12
ABK07387
ID   ABK07387 standard; RNA; 48 BP.
XX
AC   ABK07387;
XX
DT   12-MAR-2002  (first entry)
XX
```

DE Human NOGO Amberzyme substrate sequence #547.
XX
KW Human; ss; antisense therapy; cytostatic; antiinflammatory; haemostatic;
KW cerebroprotective; nootropic; neuroprotective; antiparkinsonian;
KW muscular; CD20; neurite growth inhibitor gene; NOGO; hammerhead ribozyme;
KW DNAzyme; inozyme; G-cleaver; amberzyme; zinzyme; lymphoma; leukaemia;
KW B-cell lymphoma; non-Hodgkin's lymphoma; NHL; lymphocytic leukaemia;
KW human immunodeficiency virus; HIV associated NHL; mantle-cell lymphoma;
KW MCL; immunocytoma; IMC; immune thrombocytopaenia; stroke; dementia;
KW inflammatory arthropathy; central nervous system injury;
KW cerebrovascular accident; CVA; Alzheimer's disease; multiple sclerosis;
KW chemotherapy-induced neuropathy; amyotrophic lateral sclerosis; ALS;
KW Parkinson's disease; ataxia; Huntington's disease; substrate sequence;
KW Creutzfeldt-Jakob disease; muscular dystrophy; neurodegenerative disease.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200159103-A2.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US04273.
XX
PR 11-FEB-2000; 2000US-181797P.
PR 28-FEB-2000; 2000US-185516P.
PR 06-MAR-2000; 2000US-187128P.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (BLAT/) BLATT L.
PA (MCSW/) MCSWIGGEN J.
PA (CHOW/) CHOWRIRA B M.
XX
PI Blatt L, McSwiggen J, Chowrira BM;
XX
DR WPI; 2001-607195/69.
XX
PT Nucleic acid molecules, e.g., enzymatic nucleic acids and antisense
PT constructs, which down regulate expression of a CD20 gene or neurite
PT growth inhibitor gene useful for treating, e.g., lymphoma, leukemia,
PT and central nervous system injury -
XX
PS Claim 89; Page 130; 200pp; English.
XX
CC The invention relates to a nucleic acid molecule which down regulates
CC expression of a CD20 gene and a nucleic acid molecule which down
CC regulates expression of a neurite growth inhibitor gene (NOGO).
CC The nucleic acids may be enzymatic nucleic acids (e.g. a ribozyme or a
CC DNAzyme) an Inozyme (an endolytic nucleic acid cleaving a an RNA molecule
CC possessing an NCH motif), a G-cleaver (cleaving RNA with a NYN
CC motif) pr an amberzyme (cleaving RNA with an NGN triplet), a zinzyme
CC (cleaving RNA with a YGY motif). The CD20-targetting nucleic acid is used
CC to cleave RNA of CD20 in the presence of a divalent cation that is
CC preferably Mg^2^+. Furthermore, it may be contacted with a cell to reduce
CC CD20 activity of the cell and treat a patient having a condition
CC associated with the level of CD20. The treatment may further comprise the
CC use of one or more therapies. In particular, the CD20 targetting
CC nucleic acid may be used to treat lymphoma, leukaemia, B-cell
CC lymphoma, low-grade or follicular non-Hodgkin's lymphoma (NHL), bulky
CC low-grade or follicular NHL, lymphocytic leukaemia, HIV (human
CC immunodeficiency virus) associated NHL, mantle-cell lymphoma (MCL),
CC immunocytoma (IMC), small B-cell lymphocytic lymphoma, immune
CC thrombocytopaenia, and inflammatory arthropathy. The NOGO-targetting
CC nucleic acid is used to cleave RNA of the NOGO gene in the presence of a
CC divalent cation that is preferably Mg^2^+. Furthermore, the nucleic acid
CC may be contacted with a cell to reduce NOGO activity of the cell and
CC treat a patient having a condition associated with the level of NOGO. The
CC treatment may further comprise the use of one or more therapies.
CC In particular, the NOGO-targetting nucleic acid may be used to treat
CC central nervous system (CNS) injury and cerebrovascular accident (CVA,
CC stroke), Alzheimer's disease, dementia, multiple sclerosis (MS),
CC chemotherapy-induced neuropathy, amyotrophic lateral sclerosis (ALS),
CC Parkinson's disease, ataxia, Huntington's disease, Creutzfeldt-Jakob
CC disease, muscular dystrophy, and/or other neurodegenerative disease
CC states which respond to the modulation of NOGO expression. The
CC present sequence is a substrate sequence for a nucleic acid of the
CC invention based on the human NOGO sequence.
XX
SQ Sequence 48 BP; 10 A; 13 C; 18 G; 7 U; 0 other;

Query Match 1.2%; Score 22.6; DB 23; Length 48;
Best Local Similarity 62.2%; Pred. No. 7.8e+04;
Matches 28; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

Qy 137 ggctgggggtaggggactccctacaggcacgcagccctgagaccg 181
||||:||| | ||| ||:|||: || | || | || | ||| |
Db 1 ggcugggcggaggaaacucccuucaaggacaucguccggggacug 45

RESULT 13
AAL28314
ID AAL28314 standard; DNA; 50 BP.
XX
AC AAL28314;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #1522.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 1817; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,

```
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 11 A; 16 C; 16 G; 7 T; 0 other;


  Query Match             1.2%;  Score 22.4;  DB 22;  Length 50;
  Best Local Similarity   81.2%;  Pred. No. 8.7e+04;
  Matches   26;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     349 tgcagcagacagctccctgggcaccaggcagc 380
           |||||   | ||||||||||||||| |||| |
Db      12 tgcagggaagagctccctgggcacctggcacc 43


RESULT  14
AAL34308/c
ID   AAL34308 standard; DNA; 50 BP.
XX
AC   AAL34308;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #7516.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 3554; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 3 A; 14 C; 19 G; 14 T; 0 other;


  Query Match             1.2%;  Score 22.4;  DB 22;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 8.7e+04;
  Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy     1609 catggggcaggggtgcttcctcctcaggcccagctgctcccctggagg 1656
            || |||||  |  | |||||  |  |||||| |||    ||||  ||||
Db       49 CAAGGGGCCAGACTACTTCCGCAGCAGGCCGAGCCCAACCCCAAGAGG 2


RESULT  15
AAD19606
ID   AAD19606 standard; DNA; 45 BP.
XX
AC   AAD19606;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Mutagenic PCR primer, SER-U to introduce mutation in human MKK3 gene.
XX
KW   1,3-thiazole compound; antiinflammatory; immunosuppressive; nootropic;
KW   neuroprotective; immunomodulator; antiarteriosclerotic; dermatological;
KW   antibacterial; hypertensive; p38 MAP kinase; phosphodiesterase IV;
KW   PDE IV; virucide; tumour necrosis factor alpha; TNF; Addison's disease;
KW   cytokine-mediated disease; adenosine receptor-mediated disease; diabetes;
KW   autoimmune haemolytic anaemia; Crohn's disease; psoriasis; rheumatism;
KW   spinal cord injury; multiple sclerosis; Alzheimer's disease; arthritis;
KW   Parkinson's disease; amyotrophic lateral sclerosis; ulcerative colitis;
KW   toxaemia; chronic pneumonia; pulmonary silicosis; pulmonary sarcoidosis;
KW   lung tuberculosis; Creutzfeldt-Jakob disease; arterial sclerosis; AIDS;
KW   acquired immune deficiency syndrome; systemic lupus erythematosus;
KW   cachexia; atopic dermatitis; meningitis; encephalopathy; angina pectoris;
KW   myocardial infarction; congestive heart failure; hepatitis; transplant;
KW   dialysis hypotension; diffuse intravascular coagulation syndrome;
KW   human; MKK3 gene; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200174811-A2.
XX
PD   11-OCT-2001.
XX
PF   29-MAR-2001; 2001WO-JP02629.
XX
PR   30-MAR-2000; 2000JP-0097876.
PR   02-FEB-2001; 2001JP-0027571.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Ohkawa S,  Naruo K,  Miwatashi S,  Kimura H;
XX
DR   WPI; 2001-616682/71.
XX
PT   New 1,3-thiazole compounds (I) or their salts (with position 5
PT   substituted by 4-pyridyl) for prevention and treatment of diseases
PT   including cytokine-mediated disease, adenosine receptor-mediated
PT   diseases or inflammation -
XX
PS   Example 4; Page 263; 288pp; English.
XX
CC   The present invention relates to 1,3-thiazole compounds or their
CC   salts (with position 5 substituted by 4-pyridyl group having a
CC   substituent including no aromatic group, at a position adjacent
CC   to the N of pyridyl group). The compounds of the invention are
CC   excellent inhibitors of p38 MAP kinase, phosphodiesterase IV (PDE
CC   IV) and tumour necrosis factor (TNF) alpha production. They are
CC   antagonists of adenosine receptor. They are used for the prevention
CC   and treatment of cytokine-mediated diseases, adenosine receptor-
CC   mediated diseases, inflammation, Addison's disease, autoimmune
CC   haemolytic anaemia, Crohn's disease, psoriasis, rheumatism, spinal
```

CC cord injury, multiple sclerosis, Alzheimer's disease, Parkinson's
CC disease, amyotrophic lateral sclerosis, diabetes, arthritis, toxaemias,
CC ulcerative colitis, chronic pneumonia, pulmonary silicosis, pulmonary
CC sarcoidosis, lung tuberculosis, cachexia, Creutzfeldt-Jakob disease,
CC arterial sclerosis, virus infection, atopic dermatitis, systemic lupus
CC erythematosus, meningitis, AIDS (acquired immune deficiency syndrome),
CC encephalopathy, angina pectoris, myocardial infarction, congestive
CC heart failure, hepatitis, transplant, dialysis hypotension or diffuse
CC intravascular coagulation syndrome. The present DNA sequence is PCR
CC primer, SER-U which is used for generating an active type of human
CC MKK3 gene by introducing a mutation (S189E and T193E). MKK3 gene is
CC inhibited by the compounds of the present invention.
XX
SQ Sequence 45 BP; 8 A; 9 C; 20 G; 8 T; 0 other;

Query Match 1.2%; Score 22; DB 22; Length 45;
Best Local Similarity 73.7%; Pred. No. 1e+05;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy    792 tggtgggcgaagtgctcaaggacatcgagacggcctgc 829
          |||||| ||| |||  |||||| || ||  | | ||||
Db      8 tggtggacgaggtggccaaggagatggatgccggctgc 45
```

Search completed: August 6, 2002, 22:40:31
Job time: 6285 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 20:49:36 ; Search time 88.72 Seconds
(without alignments)
5243.799 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 536926

Minimum DB seq length: 10
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	24.6	1.3	42	4	US-09-162-484-7	Sequence 7, Appli
	2	22	1.2	50	2	US-08-749-852-35	Sequence 35, Appl
	3	21.2	1.1	45	1	US-07-885-689A-7	Sequence 7, Appli
c	4	20.8	1.1	50	4	US-08-628-747-23	Sequence 23, Appl
	5	20.4	1.1	50	1	US-08-766-014-25	Sequence 25, Appl
c	6	20.2	1.1	37	2	US-08-426-599B-14	Sequence 14, Appl
	7	20.2	1.1	42	1	US-07-834-539A-50	Sequence 50, Appl
	8	20.2	1.1	42	1	US-08-053-131-67	Sequence 67, Appl
	9	20.2	1.1	42	1	US-08-645-641-67	Sequence 67, Appl
	10	20.2	1.1	42	1	US-07-853-408B-67	Sequence 67, Appl
	11	20.2	1.1	42	1	US-08-096-762-67	Sequence 67, Appl
	12	20.2	1.1	42	2	US-08-800-353-50	Sequence 50, Appl
	13	20.2	1.1	42	2	US-08-308-865-67	Sequence 67, Appl
	14	20.2	1.1	42	4	US-09-042-353-228	Sequence 228, App
	15	20.2	1.1	42	4	US-08-758-417A-76	Sequence 76, Appl
	16	20.2	1.1	42	5	PCT-US92-06185-50	Sequence 50, Appl
	17	20.2	1.1	42	5	PCT-US92-10983-67	Sequence 67, Appl
c	18	20.2	1.1	46	4	US-09-173-043-10	Sequence 10, Appl
c	19	20.2	1.1	46	4	US-09-209-525-15	Sequence 15, Appl
c	20	20	1.1	46	1	US-08-171-389-183	Sequence 183, App
c	21	20	1.1	46	1	US-08-171-389-184	Sequence 184, App
c	22	20	1.1	46	1	US-08-123-936-183	Sequence 183, App
c	23	20	1.1	46	1	US-08-123-936-184	Sequence 184, App
c	24	20	1.1	46	2	US-08-475-228A-183	Sequence 183, App
c	25	20	1.1	46	2	US-08-475-228A-184	Sequence 184, App
c	26	20	1.1	46	3	US-08-482-080A-183	Sequence 183, App
c	27	20	1.1	46	3	US-08-482-080A-184	Sequence 184, App
c	28	20	1.1	46	5	PCT-US93-12388-183	Sequence 183, App
c	29	20	1.1	46	5	PCT-US93-12388-184	Sequence 184, App
	30	19.8	1.0	44	2	US-08-883-795A-2	Sequence 2, Appli
c	31	19.8	1.0	44	2	US-08-883-795A-18	Sequence 18, Appl
c	32	19.8	1.0	46	1	US-08-171-389-182	Sequence 182, App
c	33	19.8	1.0	46	1	US-08-123-936-182	Sequence 182, App
c	34	19.8	1.0	46	2	US-08-475-228A-182	Sequence 182, App
c	35	19.8	1.0	46	3	US-08-482-080A-182	Sequence 182, App
c	36	19.8	1.0	46	5	PCT-US93-12388-182	Sequence 182, App
	37	19.6	1.0	39	1	US-07-936-110-5	Sequence 5, Appli
	38	19.6	1.0	39	1	US-08-777-918-5	Sequence 5, Appli
c	39	19.6	1.0	39	4	US-09-309-382-21	Sequence 21, Appl
	40	19.4	1.0	37	3	US-09-123-764-9	Sequence 9, Appli
c	41	19.4	1.0	43	2	US-08-531-927B-17	Sequence 17, Appl
	42	19.4	1.0	48	1	US-07-854-603-37	Sequence 37, Appl
	43	19.4	1.0	48	1	US-07-854-603-39	Sequence 39, Appl
	44	19.4	1.0	50	1	US-08-171-389-376	Sequence 376, App
c	45	19.4	1.0	50	1	US-08-171-389-455	Sequence 455, App

ALIGNMENTS

RESULT 1
US-09-162-484-7/c
; Sequence 7, Application US/09162484
; Patent No. 6248724
; GENERAL INFORMATION:
; APPLICANT: Phillips, M. Ian
; APPLICANT: Mohuczy, Dagmara
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME MRNA AND METHODS OF USE
; FILE REFERENCE: UFLA:087/UFLA087P
; CURRENT APPLICATION NUMBER: US/09/162,484
; CURRENT FILING DATE: 1998-09-25
; EARLIER APPLICATION NUMBER: 60/059,661
; EARLIER FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide
; OTHER INFORMATION: primer
US-09-162-484-7

Query Match 1.3%; Score 24.6; DB 4; Length 42;
Best Local Similarity 76.9%; Pred. No. 3.8e+03;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy    392 gccgccagcccaaacagcagcggcatgggcagcgccagc 430
          | || ||||   ||||||||||||  ||||||  ||||
Db     39 GGCGGCAGCAGCAACAGCAGCGGCAGCGGCAGCAGCAGC 1
```

RESULT 2
US-08-749-852-35
; Sequence 35, Application US/08749852
; Patent No. 5874222
; GENERAL INFORMATION:
; APPLICANT: JIRTLE, RANDY L.
; APPLICANT: DE SOUZA, ANGUS T.
; APPLICANT: HANKINS, GERALD R.
; TITLE OF INVENTION: TUMOR SUPPRESSOR
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA

```
;      COUNTRY:  U.S.A.
;      ZIP:  22201-4714
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/749,852
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WILSON, MARY J.
;      REGISTRATION NUMBER:  32,955
;      REFERENCE/DOCKET NUMBER:  1579-104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 816-4000
;      TELEFAX:  (703) 816-4100
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FRAGMENT TYPE:  linear
US-08-749-852-35


  Query Match             1.2%;  Score 22;  DB 2;  Length 50;
  Best Local Similarity   73.7%;  Pred. No. 1.4e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy     530 gagagacgggactggagtcccagtccacccgccacgcc 567
           |||   || || ||   ||||||||||||||  ||||
Db       9 GAGTCTCGTGAGTGCCTTCCCAGTCCACCCGCGGCGCC 46


RESULT   3
US-07-885-689A-7
; Sequence 7, Application US/07885689A
; Patent No. 5366876
;  GENERAL INFORMATION:
;    APPLICANT:  Cho, Joong M.
;    APPLICANT:  Lee, Tae H.
;    APPLICANT:  Chung, Hyun H.
;    APPLICANT:  Lee, Yong B.
;    APPLICANT:  Lee, Tae G.
;    APPLICANT:  Park, Young W.
;    APPLICANT:  Han, Kyu B.
;    TITLE OF INVENTION:  Method for Production of Bovine Growth
;    TITLE OF INVENTION:  Hormone Using a Synthetic Gene.
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Birch, Stewart, Kolash & Birch
;      STREET:  P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/885,689A
;      FILING DATE:  19-MAY-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  377-144P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..45
;      OTHER INFORMATION:  /label= oligonucleotide
;      OTHER INFORMATION:  /note= "U7 oligonucleotide portion of synthetic
;      OTHER INFORMATION:  BGH gene, Figure 1."
US-07-885-689A-7


  Query Match             1.1%;  Score 21.2;  DB 1;  Length 45;
  Best Local Similarity   76.5%;  Pred. No. 2.1e+04;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1691 ctctgaccccagcatttccagagcagagcctaca 1724
           ||| | |||| ||| ||||  ||||||| || ||
Db      1 CTCGGGCCCCTGCAGTTCCTCAGCAGAGTCTTCA 34


RESULT   4
US-08-628-747-23/c
; Sequence 23, Application US/08628747
; Patent No. 6169070
;  GENERAL INFORMATION:
;    APPLICANT:  Chen, Jian
;    APPLICANT:  Godowski, Paul J.
;    APPLICANT:  Hammonds, R. Glenn
;    APPLICANT:  Mark, Melanie
;    APPLICANT:  Mather, Jennie P.
;    APPLICANT:  Li, Ronghao
;    TITLE OF INVENTION:  RECEPTOR ACTIVATION BY GAS6
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  One DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/628,747
;      FILING DATE:  17-APR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/402,253
;      FILING DATE:  10-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/438,861
;      FILING DATE:  10-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lee, Wendy M.
;      REGISTRATION NUMBER:  00,000
;      REFERENCE/DOCKET NUMBER:  P929P2PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-225-1994
;      TELEFAX:  650-952-9881
;      TELEX:  910-371-7168
;  INFORMATION FOR SEQ ID NO:  23:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "oligonucleotide"
US-08-628-747-23


  Query Match             1.1%;  Score 20.8;  DB 4;  Length 50;
  Best Local Similarity   70.0%;  Pred. No. 2.6e+04;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    406 cagcagcggcatgggcagcgccagcccgggtctgagcagc 445
          |||| || ||  || | || ||| |||||| ||| | | |
Db     49 CAGCGGCAGCGGGGCCGGCCCCATCCCGGGGCTGCGGATC 10


RESULT   5
US-08-766-014-25
; Sequence 25, Application US/08766014
; Patent No. 5744312
;  GENERAL INFORMATION:
;    APPLICANT:  Mamone, Joseph A.
;    APPLICANT:  Davis, Maria
;    APPLICANT:  Sha, Dan
;    TITLE OF INVENTION:  THERMOSTABLE DNA POLYMERASE
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/766,014
;      FILING DATE:  Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  including application
;    PRIOR APPLICATION DATA:  described below:            One
;      APPLICATION NUMBER:  60/008,688
;      FILING DATE:  December 15, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  223/104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-766-014-25


  Query Match             1.1%;  Score 20.4;  DB 1;  Length 50;
  Best Local Similarity   65.2%;  Pred. No. 3.1e+04;
  Matches   30;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    279 cagatgcctggctgagacacgccagtggcctcagctgcccacacct 324
          |||  |  |   |||| || | || ||||| || | |||  |||||
Db      2 CAGTGGTATTTGTGAGCCAGGGCATTGGCCACACCAGCCACCACCT 47


RESULT   6
US-08-426-599B-14/c
; Sequence 14, Application US/08426599B
; Patent No. 5932438
;  GENERAL INFORMATION:
;    APPLICANT:  Uriach-Marsal, Juan
;    APPLICANT:  Rubio-Susan, Victor
;    APPLICANT:  Patino-Martin, Cristina
;    APPLICANT:  Kalo-Koenova, Eliza I.
;    APPLICANT:  del Moral-Juarez, Catalina
;    APPLICANT:  Faus-Santasusana, Ignacio
;    APPLICANT:  del Rio-Pericacho, Jose-Luis
;    APPLICANT:  Blade-Pique, Joan
;    TITLE OF INVENTION:  Preparation of Thaumatin Sweeteners
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg, Ernst & Kurz, P.C.
;      STREET:  555 Thirteenth Street, N.W., Suite 701 East
;      STREET:  Tower
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/426,599B
;      FILING DATE:  21-APR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saxe, Stephen A.
;      REGISTRATION NUMBER:  38,609
;      REFERENCE/DOCKET NUMBER:  1604-123A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-624-1589
;      TELEFAX:  202-783-6031
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  37 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Synthetic DNA"
US-08-426-599B-14


  Query Match             1.1%;  Score 20.2;  DB 2;  Length 37;
  Best Local Similarity   75.8%;  Pred. No. 3.2e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    841 catcaccgcagatcccatggactggagccccag 873
          |||  | |||| |||||||||  ||| |||| |
Db     35 CATGCCTGCAGGTCCCATGGAGAGGATCCCCGG 3


RESULT   7
US-07-834-539A-50
; Sequence 50, Application US/07834539A
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic Non-Human Animals Capable of
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  77
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  William M. Smith
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/834,539A
;      FILING DATE:  1992-02-05
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  14643-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-543-9600
;      TELEFAX:  415-543-5043
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-834-539A-50


  Query Match             1.1%;  Score 20.2;  DB 1;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    625 cagcagctgggcagccaaggcccctggggccag 657
          |||||| ||  || |||| |||| || | ||||
Db      1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT   8
US-08-053-131-67
; Sequence 67, Application US/08053131
; Patent No. 5661016
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 5661016-Human Animals for
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  197
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Steuart Tower, Suite 200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/053,131
;      FILING DATE:  26-APR-1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/990,860
;      FILING DATE:  16-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/810,279
;      FILING DATE:  17-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/853,408
;      FILING DATE:  18-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  14643-9-3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (primer)
US-08-053-131-67


  Query Match             1.1%;  Score 20.2;  DB 1;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    625 cagcagctgggcagccaaggcccctggggccag 657
          |||||| ||  || |||| |||| || | ||||
Db      1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT   9
US-08-645-641-67
; Sequence 67, Application US/08645641
; Patent No. 5719032
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 5719032-Human Animals for
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  150
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  William M. Smith
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/645,641
;      FILING DATE:  20-MAY-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/904,068
;      FILING DATE:  23-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  14643-000913
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (primer)
```

```
US-08-645-641-67


  Query Match             1.1%;  Score 20.2;  DB 1;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     625 cagcagctgggcagccaaggcccctggggccag 657
           |||||| ||  || |||| |||| || | ||||
Db       1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT  10
US-07-853-408B-67
; Sequence 67, Application US/07853408B
; Patent No. 5789650
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 5789650-Human Animals for
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  150
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  William M. Smith
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/853,408B
;      FILING DATE:  19920318
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  14643-9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (primer)
US-07-853-408B-67


  Query Match             1.1%;  Score 20.2;  DB 1;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     625 cagcagctgggcagccaaggcccctggggccag 657
           |||||| ||  || |||| |||| || | ||||
Db       1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT  11
US-08-096-762-67
; Sequence 67, Application US/08096762
; Patent No. 5814318
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 5814318-Human Animals for
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  210
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Steuart Tower, Suite 200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/096,762
;      FILING DATE:  22-JUL-1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,131
;      FILING DATE:  26-APR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/990,860
;      FILING DATE:  16-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/904,068
;      FILING DATE:  23-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/853,408
;      FILING DATE:  18-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/810,279
;      FILING DATE:  17-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  14643-9-4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (primer)
US-08-096-762-67


  Query Match             1.1%;  Score 20.2;  DB 1;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     625 cagcagctgggcagccaaggcccctggggccag 657
           |||||| ||  || |||| |||| || | ||||
Db       1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT  12
US-08-800-353-50
; Sequence 50, Application US/08800353
; Patent No. 5874299
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 5874299-Human Animals Capable of
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  77
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  William M. Smith
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
```

```
;       COUNTRY:  USA
;       ZIP:  94105
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/800,353
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/834,539
;       FILING DATE:  1992-02-05
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Smith, William M.
;       REGISTRATION NUMBER:  30,223
;       REFERENCE/DOCKET NUMBER:  14643-5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-543-9600
;       TELEFAX:  415-543-5043
;   INFORMATION FOR SEQ ID NO:  50:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  42 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-800-353-50


  Query Match             1.1%;  Score 20.2;  DB 2;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    625 cagcagctgggcagccaaggcccctggggccag 657
          |||||| ||  || |||| |||| || | ||||
Db      1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT  13
US-08-308-865-67
; Sequence 67, Application US/08308865
; Patent No. 5877397
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 5877397-Human Animals for
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  150
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  William M. Smith
;      STREET:  One Market Plaza, Steuart Tower, Suite 2000
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/308,865
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/145,707
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/904,068
;      FILING DATE:  23-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  14643-9-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  67:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (primer)
US-08-308-865-67


  Query Match             1.1%;  Score 20.2;  DB 2;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    625 cagcagctgggcagccaaggcccctggggccag 657
          |||||| ||  || |||| |||| || | ||||
Db      1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT  14
US-09-042-353-228
; Sequence 228, Application US/09042353
; Patent No. 6255458
;  GENERAL INFORMATION:
;    APPLICANT:  Lonberg, Nils
;    APPLICANT:  Kay, Robert M.
;    TITLE OF INVENTION:  Transgenic No. 6255458-Human Animals for
;    TITLE OF INVENTION:  Producing Heterologous Antibodies
;    NUMBER OF SEQUENCES:  421
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/042,353
;      FILING DATE:  13-MAR-1998
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/810,279
;      FILING DATE:  17-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/853,408
;      FILING DATE:  18-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/904,068
;      FILING DATE:  23-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/990,860
;      FILING DATE:  16-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/053,131
;      FILING DATE:  26-APR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/096,762
;      FILING DATE:  22-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,301
;      FILING DATE:  18-NOV-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,739
```

```
;      FILING DATE:  03-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/165,699
;      FILING DATE:  10-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/209,741
;      FILING DATE:  09-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/352,322
;      FILING DATE:  07-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/544,404
;      FILING DATE:  10-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/728,463
;      FILING DATE:  10-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US96/16433
;      FILING DATE:  10-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/758,417
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US97/21803
;      FILING DATE:  01-DEC-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Apple, Randolph T.
;      REGISTRATION NUMBER:  36,429
;      REFERENCE/DOCKET NUMBER:  014643-009040US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  228:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-042-353-228


  Query Match             1.1%;  Score 20.2;  DB 4;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     625 cagcagctgggcagccaaggcccctggggccag 657
           |||||| ||  || |||| |||| || | ||||
Db       1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


RESULT  15
US-08-758-417A-76
; Sequence 76, Application US/08758417A
; Patent No. 6300129
;   GENERAL INFORMATION:
;        APPLICANT: Lonberg, Nils
;                   Kay, Robert M.
;        TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for
;                            Producing Heterologous Antibodies
;        NUMBER OF SEQUENCES: 417
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Townsend and Townsend and Crew LLP
;             STREET: Two Embarcadero Center, Eighth Floor
;             CITY: San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94111-3834
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/758,417A
;             FILING DATE: 02-Dec-1996
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/728,463
;             FILING DATE: 10-OCT-1996
;             APPLICATION NUMBER: US 08/544,404
;             FILING DATE: 10-OCT-1995
;             APPLICATION NUMBER: US 08/352,322
;             FILING DATE: 07-DEC-1994
;             APPLICATION NUMBER: US 08/209,741
;             FILING DATE: 09-MAR-1994
;             APPLICATION NUMBER: US 08/165,699
;             FILING DATE: 10-DEC-1993
;             APPLICATION NUMBER: US 08/161,739
;             FILING DATE: 03-DEC-1993
;             APPLICATION NUMBER: US 08/155,301
;             FILING DATE: 18-NOV-1993
;             APPLICATION NUMBER: US 08/096,762
;             FILING DATE: 22-JUL-1993
;             APPLICATION NUMBER: US 08/053,131
;             FILING DATE: 26-APR-1993
;             APPLICATION NUMBER: US 07/990,860
;             FILING DATE: 16-DEC-1992
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Serafini, Andrew T.
;             REGISTRATION NUMBER: 41,303
;             REFERENCE/DOCKET NUMBER: 014643-009030US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 576-0200
;             TELEFAX: (415) 576-0300
;   INFORMATION FOR SEQ ID NO: 76:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 42 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-08-758-417A-76


  Query Match             1.1%;  Score 20.2;  DB 4;  Length 42;
  Best Local Similarity   75.8%;  Pred. No. 3.3e+04;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     625 cagcagctgggcagccaaggcccctggggccag 657
           |||||| ||  || |||| |||| || | ||||
Db       1 CAGCAGGTGCACACCCAATGCCCATGAGCCCAG 33


Search completed: August  6, 2002, 22:35:06
Job time: 6330 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 19:52:36 ; Search time 2035.51 Seconds
(without alignments)
12558.638 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 89406

Minimum DB seq length: 10
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30	1.6	47	10	R07222	R07222 yf14b07.r1
	2	26.2	1.4	40	9	AA789732	AA789732 vt78d02.r
c	3	24.8	1.3	50	9	AU104281	AU104281 AU104281
c	4	24.8	1.3	50	9	AU104284	AU104284 AU104284
c	5	23.2	1.2	50	9	AU104283	AU104283 AU104283
c	6	22.6	1.2	49	9	AA108149	AA108149 ml58h07.r
	7	21.6	1.1	46	9	AI690532	AI690532 tq12e07.x
c	8	21.6	1.1	50	9	AU104286	AU104286 AU104286
	9	21.4	1.1	47	12	AZ403867	AZ403867 1M0171A16
	10	21.4	1.1	50	9	AU104319	AU104319 AU104319
	11	21	1.1	50	9	AU107649	AU107649 AU107649
	12	21	1.1	50	9	AU107650	AU107650 AU107650
	13	21	1.1	50	9	AU107651	AU107651 AU107651
c	14	20.8	1.1	50	9	AU104063	AU104063 AU104063
	15	20.6	1.1	50	9	AU104502	AU104502 AU104502
	16	20.4	1.1	48	10	BI462075	BI462075 603205252
c	17	20.4	1.1	50	9	AU104329	AU104329 AU104329
c	18	20.4	1.1	50	9	AU105943	AU105943 AU105943
c	19	20.2	1.1	50	9	AL586305	AL586305 AL586305
c	20	20.2	1.1	50	9	AU102557	AU102557 AU102557
	21	20.2	1.1	50	9	AU104466	AU104466 AU104466
	22	20.2	1.1	50	9	AU104498	AU104498 AU104498
	23	20	1.1	44	12	AZ783407	AZ783407 2M0025P06
c	24	19.8	1.0	46	9	AA632261	AA632261 np67d09.s
	25	19.8	1.0	49	9	AA986590	AA986590 uc81g03.y
	26	19.8	1.0	50	9	AU104501	AU104501 AU104501
	27	19.8	1.0	50	9	AU104612	AU104612 AU104612
c	28	19.6	1.0	36	12	AZ807008	AZ807008 2M0069M20
	29	19.6	1.0	43	10	BE373472	BE373472 601225842
c	30	19.6	1.0	46	10	BI488502	BI488502 603021011
c	31	19.6	1.0	49	9	AI355157	AI355157 qt80g07.x
c	32	19.6	1.0	50	10	BF571906	BF571906 602078216
	33	19.4	1.0	43	9	AI570014	AI570014 tr91a09.x
c	34	19.4	1.0	47	9	AA290779	AA290779 zt22g02.r
	35	19.4	1.0	48	12	BH634120	BH634120 SALK_0447
c	36	19.4	1.0	50	9	AU102813	AU102813 AU102813
c	37	19.4	1.0	50	9	AU104330	AU104330 AU104330
c	38	19.4	1.0	50	9	AU105871	AU105871 AU105871
	39	19.4	1.0	50	10	BI256277	BI256277 602977510
c	40	19.2	1.0	42	10	BI823879	BI823879 603039045
c	41	19.2	1.0	49	9	AA677795	AA677795 zl24d03.s
c	42	19.2	1.0	49	9	AA781646	AA781646 ai50a01.s
c	43	19.2	1.0	50	9	AU104064	AU104064 AU104064
c	44	19.2	1.0	50	9	AU104491	AU104491 AU104491
c	45	19.2	1.0	50	9	AU104493	AU104493 AU104493

ALIGNMENTS

RESULT 1
R07222
LOCUS R07222 47 bp mRNA linear EST 05-APR-1995
DEFINITION yf14b07.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:126805 5' similar to gb:S45205 FLI-1 ONCOGENE (HUMAN);, mRNA sequence.
ACCESSION R07222
VERSION R07222.1 GI:759145
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 47)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1185
High qality sequence starts: 1 High qality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality
Insert Length: 1185 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .47
/organism="Homo sapiens"
/db_xref="GDB:478966"
/db_xref="taxon:9606"
/clone="IMAGE:126805"

/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 15 a 11 c 10 g 11 t
ORIGIN

Query Match 1.6%; Score 30; DB 10; Length 47;
Best Local Similarity 78.3%; Pred. No. 4.1e+04;
Matches 36; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1314 tgaactgcgacaagctgagccgctccatccgccagtcttacaagaa 1359
|||| | |||||||||||||||||| |||| | | ||| ||||
Db 1 TGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAA 46

RESULT 2
AA789732
LOCUS AA789732 40 bp mRNA linear EST 06-FEB-1998
DEFINITION vt78d02.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone IMAGE:1177251 5' similar to SW:ETS4_DROME P29775 DNA-BINDING PROTEIN D-ETS-4 ;, mRNA sequence.
ACCESSION AA789732
VERSION AA789732.1 GI:2849852
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 40)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:635099
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .40
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1177251"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead. "

BASE COUNT 10 a 10 c 11 g 9 t
ORIGIN

Query Match 1.4%; Score 26.2; DB 9; Length 40;
Best Local Similarity 79.5%; Pred. No. 2e+05;
Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1163 cacctgtggcagttcctcaaggagttgctactcaagccc 1201
|| |||||||||||||| || || || | ||||||||
Db 2 CAACTGTGGCAGTTCCTGAAAGAACTGGTGGTCAAGCCC 40

RESULT 3
AU104281/c
LOCUS AU104281 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU104281 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HEP04739, mRNA sequence.
ACCESSION AU104281
VERSION AU104281.1 GI:13553802
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
source 1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP04739"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 7 a 13 c 26 g 4 t
ORIGIN

Query Match 1.3%; Score 24.8; DB 9; Length 50;
Best Local Similarity 80.6%; Pred. No. 3.8e+05;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 457 ccacctcctgctgccccccgacacggtgtcgcggac 492
|| |||||||| |||| ||||| || |||||| | |
Db 45 CCCCCTCCTGCGGCCCGCCGACCCGCTGTCGCCGTC 10

RESULT 4
AU104284/c
LOCUS AU104284 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU104284 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HEP07220, mRNA sequence.
ACCESSION AU104284

```
VERSION     AU104284.1  GI:13553805
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP07220"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a     11 c     28 g      4 t
ORIGIN


  Query Match             1.3%;  Score 24.8;  DB 9;  Length 50;
  Best Local Similarity   80.6%;  Pred. No. 3.8e+05;
  Matches   29;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     457 ccacctcctgctgccccccgacacggtgtcgcggac 492
           || |||||||| |||| ||||| || |||||| | |
Db      44 CCCCCTCCTGCGGCCCGCCGACCCGCTGTCGCCGTC 9


RESULT   5
AU104283/c
LOCUS       AU104283                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104283 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP06519, mRNA sequence.
ACCESSION   AU104283
VERSION     AU104283.1  GI:13553804
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP06519"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a     12 c     26 g      5 t
ORIGIN


  Query Match             1.2%;  Score 23.2;  DB 9;  Length 50;
  Best Local Similarity   77.8%;  Pred. No. 7.5e+05;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     457 ccacctcctgctgccccccgacacggtgtcgcggac 492
           || |||||||| |||| ||||| || |||||  | |
Db      45 CCCCCTCCTGCGGCCCGCCGACCCGCTGTCGACGTC 10


RESULT   6
AA108149/c
LOCUS       AA108149                 49 bp    mRNA    linear   EST 03-FEB-1997
DEFINITION  ml58h07.r1 Stratagene mouse testis (#937308) Mus musculus cDNA
            clone IMAGE:516253 5' similar to gb:U29152 Mus musculus aldose
            reductase mRNA, complete cds (MOUSE);, mRNA sequence.
ACCESSION   AA108149
VERSION     AA108149.1  GI:1659201
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:310101
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /strain="Inbred CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:516253"
                     /clone_lib="Stratagene mouse testis (#937308)"
                     /sex="males"
                     /tissue_type="testis"
                     /dev_stage="10-12 week old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: testis; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Average insert size: 1.0 kb; Uni-ZAP XR Vector;
                     -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT        9 a     17 c     12 g      11 t
ORIGIN


  Query Match             1.2%;  Score 22.6;  DB 9;  Length 49;
  Best Local Similarity   68.9%;  Pred. No. 9.6e+05;
  Matches   31;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    1426 gcctggcccagggcctgaaacccgccctcaggggcctctctcctg 1470
           || ||||||||||||||  |   |  |  ||||| || | | |||
```

```
Db     45 GCTTGGCCCAGGGCCTGTCAGGAGAACCAAGGGGACTATATGCTG 1


RESULT   7
AI690532
LOCUS       AI690532                  46 bp    mRNA    linear   EST 14-DEC-1999
DEFINITION  tq12e07.x1 NCI_CGAP_Ut3 Homo sapiens cDNA clone IMAGE:2208612 3'
            similar to SW:PRPL_HUMAN P10162 SALIVARY PROLINE-RICH PROTEIN PO
            ;contains element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI690532
VERSION     AI690532.1  GI:4901834
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1032   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2208612"
                     /clone_lib="NCI_CGAP_Ut3"
                     /tissue_type="poorly-differentiated endometrial
                     adenocarcinoma, 2 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.45 kb. Life Technologies catalog #:
                     11541-018"
BASE COUNT        3 a     22 c     17 g      4 t
ORIGIN


  Query Match             1.1%;  Score 21.6;  DB 9;  Length 46;
  Best Local Similarity   68.2%;  Pred. No. 1.4e+06;
  Matches   30;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1572 ccccaggggggatatgggtcctctggggccttcgggaccatgggg 1615
           |||| |||||  |  || || ||||| || ||    || |||||
Db      3 CCCCGGGGGGGAAACGGGCCCCTGGGCCCCTCCCCCCCTTGGGG 46


RESULT   8
AU104286/c
LOCUS       AU104286                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104286 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP15925, mRNA sequence.
ACCESSION   AU104286
VERSION     AU104286.1  GI:13553807
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP15925"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a     11 c     25 g      7 t
ORIGIN


  Query Match             1.1%;  Score 21.6;  DB 9;  Length 50;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+06;
  Matches   27;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    457 ccacctcctgctgccccccgacacggtgtcgcggac 492
          || |||||||| ||||   |||||  |||||| | |
Db     45 CCCCCTCCTGCGGCCCGAAGACACACTGTCGCCGTC 10


RESULT   9
AZ403867
LOCUS       AZ403867                  47 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0171A16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0171A16 R, DNA sequence.
ACCESSION   AZ403867
VERSION     AZ403867.1  GI:10527880
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0171  row: A  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 47.
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0171A16"
```

```
                /clone_lib="Mouse 10kb plasmid UUGC1M library"
                /sex="Male"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /note="Vector: PWD42nv; Purified genomic DNA from M.
                musculus C57BL/6J (male) was obtained  from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a
                0.005 inch orifice at constant velocity. The sheared DNA
                was blunt end-repaired with T4 DNA polymerase and T4
                polynucleotide kinase. Adaptor oligonucleotides were
                ligated to the blunt ends in high  molar excess. The
                adaptored DNA was purified and size-selected for a 9.5 to
                10.5 kb range using preparative agarose gel
                electrophoresis. Vector DNA was prepared from a derivative
                of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adaptored mouse DNA was annealed to
                adaptored vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
BASE COUNT        8 a     13 c     15 g     11 t
ORIGIN


  Query Match             1.1%;  Score 21.4;  DB 12;  Length 47;
  Best Local Similarity   71.8%;  Pred. No. 1.6e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1020 tgcacgcccacctggacatctggaagtcagcggcctgga 1058
           | || |||||||||| | ||| | ||  || ||| || |
Db      9 TCCATGCCCACCTGGTCTTCTTGGAGGGAGAGGCATGCA 47


RESULT  10
AU104319
LOCUS       AU104319                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104319 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP05158, mRNA sequence.
ACCESSION   AU104319
VERSION     AU104319.1  GI:13553840
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP05158"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       11 a     11 c     18 g     10 t
ORIGIN


  Query Match             1.1%;  Score 21.4;  DB 9;  Length 50;
  Best Local Similarity   66.0%;  Pred. No. 1.6e+06;
  Matches   31;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    778 ggtgcagtccatggtggtgggcgaagtgctcaaggacatcgagacgg 824
          ||||  |||||| ||||   ||| | |  | |||| ||| | | | |
Db      4 GGTGGTGTCCATCGTGGGCAGCGGACTAATAAAGGCCATGGCGCCAG 50


RESULT  11
AU107649
LOCUS       AU107649                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107649 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI02001, mRNA sequence.
ACCESSION   AU107649
VERSION     AU107649.1  GI:13557170
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI02001"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       17 a     11 c     15 g      7 t
ORIGIN


  Query Match             1.1%;  Score 21;  DB 9;  Length 50;
  Best Local Similarity   82.8%;  Pred. No. 1.9e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    862 ctggagccccagcaatgtgcagaagtggc 890
          | ||||||| |||| |||| ||||| |||
Db     21 CCGGAGCCCGAGCAGTGTGAAGAAGAGGC 49


RESULT  12
AU107650
LOCUS       AU107650                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107650 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI03379, mRNA sequence.
ACCESSION   AU107650
VERSION     AU107650.1  GI:13557171
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
```

```
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI03379"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       17 a     13 c     14 g      6 t
ORIGIN


  Query Match             1.1%;  Score 21;  DB 9;  Length 50;
  Best Local Similarity   82.8%;  Pred. No. 1.9e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy     862 ctggagccccagcaatgtgcagaagtggc 890
           | ||||||| |||| |||| ||||| |||
Db      16 CCGGAGCCCGAGCAGTGTGAAGAAGAGGC 44


RESULT  13
AU107651
LOCUS       AU107651                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107651 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI07783, mRNA sequence.
ACCESSION   AU107651
VERSION     AU107651.1  GI:13557172
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI07783"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT       18 a     12 c     13 g      7 t
ORIGIN


  Query Match             1.1%;  Score 21;  DB 9;  Length 50;
  Best Local Similarity   82.8%;  Pred. No. 1.9e+06;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy     862 ctggagccccagcaatgtgcagaagtggc 890
           | ||||||| |||| |||| ||||| |||
Db      21 CCGGAGCCCGAGCAGTGTGAAGAAGAGGC 49


RESULT  14
AU104063/c
LOCUS       AU104063                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104063 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS01643, mRNA sequence.
ACCESSION   AU104063
VERSION     AU104063.1  GI:13553584
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS01643"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        7 a     15 c     15 g     13 t
ORIGIN


  Query Match             1.1%;  Score 20.8;  DB 9;  Length 50;
  Best Local Similarity   78.1%;  Pred. No. 2.1e+06;
  Matches   25;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy    1330 gagccgctccatccgccagtcttacaagaagg 1361
           ||||||| || |||||||| | | ||||||
Db      44 GAGCCGCCGCAGCCGCCAGTTCTGCTAGAAGG 13


RESULT  15
AU104502
LOCUS       AU104502                  50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU104502 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            ZRV6C934, mRNA sequence.
ACCESSION   AU104502
VERSION     AU104502.1  GI:13554023
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
```

Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano ,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

```
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="ZRV6C934"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT        9 a      9 c     25 g      7 t
ORIGIN
```

Query Match 1.1%; Score 20.6; DB 9; Length 50;
Best Local Similarity 74.3%; Pred. No. 2.3e+06;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy     501 agaaggcggcagcgggggcagtgggtctcgagaga 535
           ||| |||||| |||| ||  ||||   || |||||
Db      12 AGATGGCGGCGGCGGCGGTGGTGGAGTTCCAGAGA 46
```

Search completed: August 6, 2002, 21:35:03
Job time: 6147 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 21:01:12 ; Search time 3320.87 Seconds
(without alignments)
12339.085 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 10010710

Minimum DB seq length: 10
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US100_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US101_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	34	1.8	34	22	US-09-570-593-11	Sequence 11, Appl
	2	31	1.6	31	22	US-09-570-593-10	Sequence 10, Appl
c	3	24.4	1.3	50	29	US-09-755-374A-3340	Sequence 3340, Ap
	4	24.2	1.3	50	29	US-09-755-374A-26208	Sequence 26208, A
	5	23.8	1.3	50	29	US-09-755-374A-2825	Sequence 2825, Ap
	6	23.6	1.2	48	37	US-10-017-974-35747	Sequence 35747, A
	7	23.2	1.2	47	18	US-09-422-978-175	Sequence 175, App
	8	23	1.2	23	28	US-09-718-159-5	Sequence 5, Appli
	9	23	1.2	23	28	US-09-718-159-9	Sequence 9, Appli
	10	23	1.2	23	28	US-09-718-159-10	Sequence 10, Appl
	11	23	1.2	23	28	US-09-718-159-11	Sequence 11, Appl
	12	23	1.2	50	29	US-09-755-374A-9716	Sequence 9716, Ap
	13	23	1.2	50	64	US-60-253-652-21580	Sequence 21580, A
	14	22.8	1.2	50	29	US-09-755-374A-28060	Sequence 28060, A
c	15	22.6	1.2	38	8	US-08-472-801-161	Sequence 161, App
c	16	22.6	1.2	38	10	US-08-668-235-161	Sequence 161, App
	17	22.6	1.2	48	30	US-09-780-533A-6295	Sequence 6295, Ap
	18	22.4	1.2	48	33	US-09-864-785-3315	Sequence 3315, Ap
	19	22.2	1.2	41	35	US-09-953-198-123	Sequence 123, App
c	20	22.2	1.2	44	8	US-08-472-801-999	Sequence 999, App
c	21	22.2	1.2	44	10	US-08-668-235-1000	Sequence 1000, Ap
	22	22.2	1.2	48	37	US-10-017-974-22699	Sequence 22699, A
	23	22	1.2	22	14	US-09-064-645-23	Sequence 23, Appl
	24	22	1.2	22	22	US-09-567-970-23	Sequence 23, Appl
	25	22	1.2	45	35	US-09-959-356A-7	Sequence 7, Appli
c	26	22	1.2	45	35	US-09-959-356A-8	Sequence 8, Appli
	27	22	1.2	48	33	US-09-864-785-3326	Sequence 3326, Ap
	28	22	1.2	48	37	US-10-017-974-34757	Sequence 34757, A
c	29	22	1.2	50	29	US-09-755-374A-17016	Sequence 17016, A
c	30	21.8	1.2	49	64	US-60-253-652-9804	Sequence 9804, Ap
c	31	21.8	1.2	50	29	US-09-755-374A-24550	Sequence 24550, A

	32	21.6	1.1	31	31	US-09-801-274-1269 Sequence 1269, Ap
c	33	21.6	1.1	47	60	US-60-216-745-4273 Sequence 4273, Ap
	34	21.6	1.1	48	35	US-09-930-423-3875 Sequence 3875, Ap
c	35	21.6	1.1	50	29	US-09-755-374A-23624 Sequence 23624, A
c	36	21.6	1.1	50	29	US-09-755-374A-23626 Sequence 23626, A
	37	21.4	1.1	48	32	US-09-848-754A-7792 Sequence 7792, Ap
	38	21.4	1.1	48	33	US-09-864-785-3298 Sequence 3298, Ap
c	39	21.2	1.1	36	16	US-09-214-913-51 Sequence 51, Appl
c	40	21.2	1.1	36	23	US-09-612-314-64 Sequence 64, Appl
	41	21.2	1.1	36	30	US-09-762-278-11 Sequence 11, Appl
	42	21	1.1	21	28	US-09-718-159-3 Sequence 3, Appli
c	43	21	1.1	37	22	US-09-593-580-45 Sequence 45, Appl
c	44	21	1.1	37	22	US-09-593-580B-45 Sequence 45, Appl
	45	21	1.1	39	30	US-09-771-372-8 Sequence 8, Appli

ALIGNMENTS

```
RESULT   1
US-09-570-593-11/c
; Sequence 11, Application US/09570593
; GENERAL INFORMATION:
;  APPLICANT:  Kaufmann, Joerg
;  APPLICANT:  Xin, Hong
;  APPLICANT:  Harrowe, Greg
;  TITLE OF INVENTION:  EXPRESSION OF ETS-DOMAIN PROTEINS IN
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE:  2300-1556
;  CURRENT APPLICATION NUMBER: US/09/570,593
;  CURRENT FILING DATE:  2000-05-12
;  PRIOR APPLICATION NUMBER:  60/134,112
;  PRIOR FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: GSEF-specific primer
US-09-570-593-11


  Query Match             1.8%;  Score 34;  DB 22;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   34;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1850 ccaggagtgccccgggaatggataataaagata 1883
          ||||||||||||||||||||||||||||||||||
Db     34 CCAGGAGTGCCCCCGGGAATGGATAATAAAGATA 1


RESULT   2
US-09-570-593-10
; Sequence 10, Application US/09570593
; GENERAL INFORMATION:
;  APPLICANT:  Kaufmann, Joerg
;  APPLICANT:  Xin, Hong
;  APPLICANT:  Harrowe, Greg
;  TITLE OF INVENTION:  EXPRESSION OF ETS-DOMAIN PROTEINS IN
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE:  2300-1556
;  CURRENT APPLICATION NUMBER: US/09/570,593
;  CURRENT FILING DATE:  2000-05-12
;  PRIOR APPLICATION NUMBER:  60/134,112
;  PRIOR FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: PArtificial Sequence
;   FEATURE:
;   OTHER INFORMATION: GSEF-specific primer
US-09-570-593-10


  Query Match             1.6%;  Score 31;  DB 22;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 5e+04;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1552 cagggaggggcaaccaactgccccaggggga 1582
          |||||||||||||||||||||||||||||||
Db      1 cagggaggggcaaccaactgccccaggggga 31


RESULT   3
US-09-755-374A-3340/c
; Sequence 3340, Application US/09755374A
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and M
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 15966-611
;  CURRENT APPLICATION NUMBER: US/09/755,374A
;  CURRENT FILING DATE:  2001-01-08
;  PRIOR APPLICATION NUMBER: 60/174962
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 3340
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 2 of 2 allelic variants (3339 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (25)...(26)
;   OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg44930423
US-09-755-374A-3340


  Query Match             1.3%;  Score 24.4;  DB 29;  Length 50;
  Best Local Similarity   68.0%;  Pred. No. 9.1e+05;
  Matches   34;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    399 gcccaaacagcagcggcatgggcagcgccagcccgggtctgagcagcgta 448
          ||| ||| |||||    || |||| | |||| ||  |  ||||||| | |
Db     50 GCCTAAAAAGCAGACAGATAGGCATCCCCAGGCCCTGAGTGAGCAGAGGA 1


RESULT   4
US-09-755-374A-26208
; Sequence 26208, Application US/09755374A
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and M
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 15966-611
;  CURRENT APPLICATION NUMBER: US/09/755,374A
;  CURRENT FILING DATE:  2001-01-08
;  PRIOR APPLICATION NUMBER: 60/174962
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 26208
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 2 of 2 allelic variants (26207 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (25)...(26)
;   OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg44009035
US-09-755-374A-26208


  Query Match             1.3%;  Score 24.2;  DB 29;  Length 50;
  Best Local Similarity   71.1%;  Pred. No. 9.9e+05;
  Matches   32;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    968 gccatgtcggaggagcagttccgccagcgctcgcccctgggtggg 1012
          | ||||  ||  | || || | ||||| ||| || |||||| |||
Db      5 ggcatgaggggagggctgtgctgccagggctagcacctgggaggg 49


RESULT   5
US-09-755-374A-2825
; Sequence 2825, Application US/09755374A
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and Meth
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 15966-611
;  CURRENT APPLICATION NUMBER: US/09/755,374A
;  CURRENT FILING DATE:  2001-01-08
;  PRIOR APPLICATION NUMBER: 60/174962
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 2825
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 1 of 2 allelic variants (2826 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (25)...(26)
;   OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg43969816
US-09-755-374A-2825


  Query Match             1.3%;  Score 23.8;  DB 29;  Length 50;
  Best Local Similarity   72.1%;  Pred. No. 1.2e+06;
  Matches   31;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1124 gacagcgaggtggactcatcatgctccgggcagcccatccacc 1166
          | |||   ||||| |||  ||| ||||  ||  ||||||||||
Db      5 gccaggctggtgggctccgcattctccaagccccccatccacc 47


RESULT   6
US-10-017-974-35747
; Sequence 35747, Application US/10017974
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Blatt, Lawrence
;  APPLICANT:  McSwiggen, James
;  TITLE OF INVENTION: Nucleic Acid-Based Treatment of Diseases or Conditions Related to
;  TITLE OF INVENTION:  Nile Virus Infection
;  FILE REFERENCE: MBHB00,1109-A (400/037)
;  CURRENT APPLICATION NUMBER: US/10/017,974
;  CURRENT FILING DATE:  2001-12-10
;  NUMBER OF SEQ ID NOS: 37080
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35747
;   LENGTH: 48
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-10-017-974-35747


  Query Match             1.2%;  Score 23.6;  DB 37;  Length 48;
  Best Local Similarity   63.0%;  Pred. No. 1.3e+06;
  Matches   29;  Conservative    3;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    137 ggctgggggtaggggactccctacaggcacgcagccctgagaccgc 182
          |||:||||| |||  ||:|||: || | ||   | || |  || ||
Db      1 ggcuggggggaggaaacucccuucaaggacaucguccgggaacggc 46


RESULT   7
US-09-422-978-175
; Sequence 175, Application US/09422978
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 175
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-13320-352 : polymorphic base A or G
US-09-422-978-175


  Query Match             1.2%;  Score 23.2;  DB 18;  Length 47;
  Best Local Similarity   67.4%;  Pred. No. 1.5e+06;
  Matches   31;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1484 agccaggccctgagatgggggaaaacgggcagtgtgctctgctgct 1529
          || || ||||||        || |: | ||| ||||||||||||| |
Db      1 aggcaagccctgtaccctcgggargccggctgtgtgctctgctgtt 46


RESULT   8
US-09-718-159-5
; Sequence 5, Application US/09718159
; GENERAL INFORMATION:
;  APPLICANT:  Harkins, Richard
;  APPLICANT:  Lin, Richard
;  APPLICANT:  Luke, May
;  APPLICANT:  Monteclaro, Felipe
;  APPLICANT:  Parkes, Deborah
;  APPLICANT:  Parry, Gordon
;  APPLICANT:  Steinbrecher, Renate
;  APPLICANT:  Van Heuit, Pamela Toy
;  APPLICANT:  Xuan, Jian-ai
;  TITLE OF INVENTION:  DNA Encoding a Novel ETS Polypeptide
;  FILE REFERENCE:  51790AUSM1
```

```
;  CURRENT APPLICATION NUMBER: US/09/718,159
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER:  60/168,182
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: probe
US-09-718-159-5


  Query Match             1.2%;  Score 23;  DB 28;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1783 cctgcctcccattctgcaccaca 1805
          |||||||||||||||||||||||
Db      1 cctgcctcccattctgcaccaca 23


RESULT   9
US-09-718-159-9
; Sequence 9, Application US/09718159
; GENERAL INFORMATION:
;  APPLICANT:  Harkins, Richard
;  APPLICANT:  Lin, Richard
;  APPLICANT:  Luke, May
;  APPLICANT:  Monteclaro, Felipe
;  APPLICANT:  Parkes, Deborah
;  APPLICANT:  Parry, Gordon
;  APPLICANT:  Steinbrecher, Renate
;  APPLICANT:  Van Heuit, Pamela Toy
;  APPLICANT:  Xuan, Jian-ai
;  TITLE OF INVENTION:  DNA Encoding a Novel ETS Polypeptide
;  FILE REFERENCE:  51790AUSM1
;  CURRENT APPLICATION NUMBER: US/09/718,159
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER:  60/168,182
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 23
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: antisense
US-09-718-159-9


  Query Match             1.2%;  Score 23;  DB 28;  Length 23;
  Best Local Similarity   73.9%;  Pred. No. 1.6e+06;
  Matches   17;  Conservative    6;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1161 tccacctgtggcagttcctcaag 1183
          :|||||:|:||||||::||:||||
Db      1 uccaccuguggcaguuccucaag 23


RESULT  10
US-09-718-159-10
; Sequence 10, Application US/09718159
; GENERAL INFORMATION:
;  APPLICANT:  Harkins, Richard
;  APPLICANT:  Lin, Richard
;  APPLICANT:  Luke, May
;  APPLICANT:  Monteclaro, Felipe
;  APPLICANT:  Parkes, Deborah
;  APPLICANT:  Parry, Gordon
;  APPLICANT:  Steinbrecher, Renate
;  APPLICANT:  Van Heuit, Pamela Toy
;  APPLICANT:  Xuan, Jian-ai
;  TITLE OF INVENTION:  DNA Encoding a Novel ETS Polypeptide
;  FILE REFERENCE:  51790AUSM1
;  CURRENT APPLICATION NUMBER: US/09/718,159
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER:  60/168,182
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 23
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: antisense
US-09-718-159-10


  Query Match             1.2%;  Score 23;  DB 28;  Length 23;
  Best Local Similarity   91.3%;  Pred. No. 1.6e+06;
  Matches   21;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1734 tgactcgacaaaggccacaggca 1756
          :|||:||||||||||||||||||
Db      1 ugacucgacaaaggccacaggca 23


RESULT  11
US-09-718-159-11
; Sequence 11, Application US/09718159
; GENERAL INFORMATION:
;  APPLICANT:  Harkins, Richard
;  APPLICANT:  Lin, Richard
;  APPLICANT:  Luke, May
;  APPLICANT:  Monteclaro, Felipe
;  APPLICANT:  Parkes, Deborah
;  APPLICANT:  Parry, Gordon
;  APPLICANT:  Steinbrecher, Renate
;  APPLICANT:  Van Heuit, Pamela Toy
;  APPLICANT:  Xuan, Jian-ai
;  TITLE OF INVENTION:  DNA Encoding a Novel ETS Polypeptide
;  FILE REFERENCE:  51790AUSM1
;  CURRENT APPLICATION NUMBER: US/09/718,159
;  CURRENT FILING DATE:  2000-11-21
;  PRIOR APPLICATION NUMBER:  60/168,182
;  PRIOR FILING DATE:  1999-11-30
;  NUMBER OF SEQ ID NOS:  11
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 23
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: antisense
US-09-718-159-11


  Query Match             1.2%;  Score 23;  DB 28;  Length 23;
  Best Local Similarity   82.6%;  Pred. No. 1.6e+06;
  Matches   19;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1699 ccagcatttccagagcagagcct 1721
          |||||||:::|||||||||||||:
Db      1 ccagcauuuccagagcagagccu 23


RESULT  12
US-09-755-374A-9716
; Sequence 9716, Application US/09755374A
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin
```

```
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and Meth
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 15966-611
;  CURRENT APPLICATION NUMBER: US/09/755,374A
;  CURRENT FILING DATE:  2001-01-08
;  PRIOR APPLICATION NUMBER: 60/174962
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 9716
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 2 of 2 allelic variants (9715 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (25)...(26)
;   OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg43952512
US-09-755-374A-9716


  Query Match             1.2%;  Score 23;  DB 29;  Length 50;
  Best Local Similarity   74.4%;  Pred. No. 1.7e+06;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    629 agctgggcagccaaggcccctggggccagcagtcgggag 667
          |||||  | | | ||| || ||||||||||| | | ||
Db      1 agctgcactggccaggaccatggggccagcaggctgcag 39


RESULT  13
US-60-253-652-21580
; Sequence 21580, Application US/60253652
; GENERAL INFORMATION:
;  APPLICANT: Glenn, Matthew
;  APPLICANT:  Grigor, Murray R
;  TITLE OF INVENTION: Compositions isolated from bovine
;  TITLE OF INVENTION:  tissues and methods for their use.
;  FILE REFERENCE: 1055P2
;  CURRENT APPLICATION NUMBER: US/60/253,652
;  CURRENT FILING DATE:  2000-11-28
;  NUMBER OF SEQ ID NOS: 29954
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21580
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Bovine
US-60-253-652-21580


  Query Match             1.2%;  Score 23;  DB 64;  Length 50;
  Best Local Similarity   74.4%;  Pred. No. 1.7e+06;
  Matches   29;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    383 acagacacagccgccagcccaaacagcagcggcatgggc 421
          || || |||| ||| |  ||| || || ||| |||||||
Db     12 actgagacaggcgctacaccatactgcggcgccatgggc 50


RESULT  14
US-09-755-374A-28060
; Sequence 28060, Application US/09755374A
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide Polymorphisms and Meth
;  TITLE OF INVENTION:  Use Thereof
;  FILE REFERENCE: 15966-611
;  CURRENT APPLICATION NUMBER: US/09/755,374A
;  CURRENT FILING DATE:  2001-01-08
;  PRIOR APPLICATION NUMBER: 60/174962
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 28742
; SEQ ID NO 28060
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 2 of 2 allelic variants (28059 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (25)...(26)
;   OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg44009326
US-09-755-374A-28060


  Query Match             1.2%;  Score 22.8;  DB 29;  Length 50;
  Best Local Similarity   66.0%;  Pred. No. 1.8e+06;
  Matches   33;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy    485 tcgcggacaggcttggagaaggcggcagcgggggcagtgggtctcgagag 534
          |||||| |||||| || |     ||| ||||| |  || |||  ||  ||
Db      1 tcgcggccaggctgggcgcgcagggcggcgggtggtgttggttacgtcag 50


RESULT  15
US-08-472-801-161/c
; Sequence 161, Application US/08472801
; GENERAL INFORMATION:
;  APPLICANT: Hesed 2
;  APPLICANT: Smith, Larry J.
;  TITLE OF INVENTION: Method and Compositions for Cellular
;  TITLE OF INVENTION: Reprogramming
;  FILE REFERENCE: Hesed 2
;  CURRENT APPLICATION NUMBER: US/08/472,801
;  CURRENT FILING DATE: 1995-06-07
;  NUMBER OF SEQ ID NOS: 3601
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 161
;   LENGTH: 38
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-08-472-801-161


  Query Match             1.2%;  Score 22.6;  DB 8;  Length 38;
  Best Local Similarity   75.7%;  Pred. No. 1.9e+06;
  Matches   28;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    958 ggagctgtgcgccatgtcggaggagcagttccgccag 994
          |||| || |||||| |  ||||||||||  ||  |||
Db     37 GGAGGTGCGCGCCAAGCTGGAGGAGCAGGCCCAGCAG 1


Search completed: August  6, 2002, 23:36:06
Job time: 9294 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: August 6, 2002, 21:35:07 ; Search time 378.9 Seconds
(without alignments)
10877.386 Million cell updates/sec

Title: US-09-841-963A-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1362792 seqs, 1088025756 residues

Total number of hits satisfying chosen parameters: 646124

Minimum DB seq length: 10
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq2:*
8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	21.8	1.2	41	8	US-60-384-980-216	Sequence 216, App
	2	21.6	1.1	48	5	US-09-745-237A-3875	Sequence 3875, Ap
c	3	21.2	1.1	50	5	US-09-980-217-51	Sequence 51, Appl
	4	21	1.1	41	1	PCT-US02-12063-268	Sequence 268, App
	5	21	1.1	41	7	US-10-126-022-268	Sequence 268, App
	6	20.8	1.1	48	7	US-10-156-306-7580	Sequence 7580, Ap
c	7	20.8	1.1	50	5	US-09-718-321A-1048	Sequence 1048, Ap
c	8	20.6	1.1	49	5	US-09-978-917A-25	Sequence 25, Appl
c	9	20.6	1.1	49	5	US-09-997-623-25	Sequence 25, Appl
	10	20.4	1.1	48	5	US-09-745-237A-3833	Sequence 3833, Ap
c	11	20.2	1.1	41	8	US-60-384-980-69	Sequence 69, Appl
c	12	20.2	1.1	50	7	US-10-099-926-1257	Sequence 1257, Ap
	13	20	1.1	48	5	US-09-745-237A-3872	Sequence 3872, Ap
	14	20	1.1	48	5	US-09-745-237A-3877	Sequence 3877, Ap
c	15	19.8	1.0	44	5	US-09-913-238-45	Sequence 45, Appl
c	16	19.8	1.0	47	7	US-10-170-097-1121	Sequence 1121, Ap
	17	19.8	1.0	48	5	US-09-745-237A-4024	Sequence 4024, Ap
	18	19.8	1.0	48	5	US-09-745-237A-4051	Sequence 4051, Ap
	19	19.8	1.0	48	7	US-10-156-306-3945	Sequence 3945, Ap
	20	19.8	1.0	48	7	US-10-156-306-4297	Sequence 4297, Ap
	21	19.8	1.0	48	7	US-10-156-306-7660	Sequence 7660, Ap
c	22	19.8	1.0	50	5	US-09-718-321A-1047	Sequence 1047, Ap
c	23	19.6	1.0	43	7	US-10-139-480-18	Sequence 18, Appl
	24	19.6	1.0	48	5	US-09-745-237A-3850	Sequence 3850, Ap
c	25	19.6	1.0	50	5	US-09-718-321A-228	Sequence 228, App
	26	19.4	1.0	40	7	US-10-068-965-5	Sequence 5, Appli
	27	19.4	1.0	41	1	PCT-US02-12063-402	Sequence 402, App
	28	19.4	1.0	41	7	US-10-126-022-402	Sequence 402, App
	29	19.4	1.0	48	5	US-09-745-237A-4248	Sequence 4248, Ap
	30	19.4	1.0	48	7	US-10-156-306-7370	Sequence 7370, Ap
	31	19.2	1.0	40	7	US-10-109-349A-102	Sequence 102, App
	32	19.2	1.0	48	5	US-09-745-237A-4318	Sequence 4318, Ap
c	33	19.2	1.0	50	1	PCT-US02-12405-476	Sequence 476, App
c	34	19.2	1.0	50	7	US-10-144-094-33	Sequence 33, Appl
c	35	19.2	1.0	50	7	US-10-045-674-476	Sequence 476, App
	36	19	1.0	43	5	US-09-747-774A-80	Sequence 80, Appl
	37	19	1.0	45	5	US-09-380-447A-246	Sequence 246, App
c	38	19	1.0	47	7	US-10-170-097-971	Sequence 971, App
	39	19	1.0	48	5	US-09-745-237A-3916	Sequence 3916, Ap
	40	19	1.0	48	5	US-09-745-237A-4015	Sequence 4015, Ap
	41	19	1.0	48	7	US-10-156-306-4143	Sequence 4143, Ap
	42	19	1.0	48	7	US-10-156-306-7410	Sequence 7410, Ap
	43	19	1.0	48	7	US-10-156-306-7653	Sequence 7653, Ap
c	44	19	1.0	50	5	US-09-718-321A-1013	Sequence 1013, Ap
c	45	18.8	1.0	41	8	US-60-384-980-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-60-384-980-216/c
; Sequence 216, Application US/60384980
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0211 PSP
; CURRENT APPLICATION NUMBER: US/60/384,980
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 1219
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 216
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-384-980-216

Query Match 1.2%; Score 21.8; DB 8; Length 41;
Best Local Similarity 70.7%; Pred. No. 3.1e+05;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy   1233 acaaggagaagggcatcttcaaaattgaggactcagcccag 1273
          | | || |||| | || || ||| ||||| ||||  |||
Db     41 AAAGGGGGAAGAGAATGTTAAAATCTGAGGTCTCAGGGCAG 1
```

RESULT 2
US-09-745-237A-3875
; Sequence 3875, Application US/09745237A
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
; FILE REFERENCE: 400/007 (MBHB00-918-A)
; CURRENT APPLICATION NUMBER: US/09/745,237A
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 4550
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3875
; LENGTH: 48
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-09-745-237A-3875

Query Match 1.1%; Score 21.6; DB 5; Length 48;
Best Local Similarity 63.6%; Pred. No. 3.4e+05;
Matches 28; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

```
Qy    141 gggggtaggggactccctacaggcacgcagccctgagaccgcca 184
          ||||| |||  ||:|||: || | ||   | || | | || |||
Db      5 gggggggaggaaacucccuucaaggacaucguccggggccccccca 48
```

RESULT 3
US-09-980-217-51/c
; Sequence 51, Application US/09980217
; GENERAL INFORMATION:
; APPLICANT: Biotica Technology Limited
; APPLICANT: Leadlay, Peter F
; APPLICANT: Staunton, James
; APPLICANT: Oliynyk, Marko
; TITLE OF INVENTION: Polyketides and their synthesis
; FILE REFERENCE: IS/BP5858469
; CURRENT APPLICATION NUMBER: US/09/980,217
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: PCT/GB00/02072
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: GB 9912563.5
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-980-217-51

Query Match 1.1%; Score 21.2; DB 5; Length 50;
Best Local Similarity 76.5%; Pred. No. 4.1e+05;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy     27 gcactctgccgtgtccacactgccccacagaccc 60
          | |||  |||||||| |||||| || || | |||
Db     43 GGACTGCGCCGTGTCGACACTGACCGACGGGCCC 10
```

RESULT 4
PCT-US02-12063-268
; Sequence 268, Application PC/TUS0212063
; GENERAL INFORMATION:
; APPLICANT: SCHERING CORPORATION
; APPLICANT: GENOME THERAPEUTICS CORPORATION
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 2976-4039PC2
; CURRENT APPLICATION NUMBER: PCT/US02/12063
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 09/834,597
; PRIOR FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 268
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: containing SNP
PCT-US02-12063-268

Query Match 1.1%; Score 21; DB 1; Length 41;
Best Local Similarity 82.8%; Pred. No. 4.4e+05;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy    624 acagcagctgggcagccaaggcccctggg 652
          |||||  |||| |||||  ||||||||||
Db      6 acagccactggacagccctggcccctggg 34
```

RESULT 5
US-10-126-022-268
; Sequence 268, Application US/10126022
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 2976-4039US2
; CURRENT APPLICATION NUMBER: US/10/126,022
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/834,597
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/548,797
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 268
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: containing SNP
US-10-126-022-268

Query Match 1.1%; Score 21; DB 7; Length 41;
Best Local Similarity 82.8%; Pred. No. 4.4e+05;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy    624 acagcagctgggcagccaaggcccctggg 652
          |||||  |||| |||||  ||||||||||
Db      6 acagccactggacagccctggcccctggg 34
```

RESULT 6
US-10-156-306-7580
; Sequence 7580, Application US/10156306
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Rel
; TITLE OF INVENTION: Levels of IKK-Gamma and PKR
; FILE REFERENCE: MBHB01-664-A (400/050)
; CURRENT APPLICATION NUMBER: US/10/156,306
; CURRENT FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 8013
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7580
; LENGTH: 48
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid
US-10-156-306-7580

Query Match 1.1%; Score 20.8; DB 7; Length 48;
Best Local Similarity 60.4%; Pred. No. 4.8e+05;
Matches 29; Conservative 2; Mismatches 17; Indels 0; Gaps 0;

```
Qy    137 ggctgggggtaggggactccctacaggcacgcagccctgagaccgcca 184
          ||| ||||| |||  ||:|||: || | ||   | || | |   | ||
Db      1 ggcaggggggaggaaacucccuucaaggacaucguccgggggggggca 48
```

```
RESULT   7
US-09-718-321A-1048/c
; Sequence 1048, Application US/09718321A
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard A.
;  APPLICANT:  Leach, Martin D.
;  TITLE OF INVENTION: NUCLEIC ACIDS CONTAINING SINGLE NUCLEIC ACID POLYMORPHISMS AND ME
;  TITLE OF INVENTION:  USE THEREOF
;  FILE REFERENCE: 15966-599
;  CURRENT APPLICATION NUMBER: US/09/718,321A
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: 60/163,783
;  PRIOR FILING DATE: 1999-11-24
;  NUMBER OF SEQ ID NOS: 1468
;  SOFTWARE: CuraGen Patent Formatter Version 0.9
; SEQ ID NO 1048
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: single nucleotide polymorphism
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (25)...(26)
;   OTHER INFORMATION: Nucleotide deleted between bases 25 and 26
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg43955639
US-09-718-321A-1048


  Query Match             1.1%;  Score 20.8;  DB 5;  Length 50;
  Best Local Similarity   64.6%;  Pred. No. 4.8e+05;
  Matches   31;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy   1640 agctgctcccctggaggacagagggagacagggctgctccccaacacc 1687
          ||| ||| |||||||||||  |      ||||    | ||||||  |||
Db     49 AGCCGCTACCCTGGAGGAGGGGAACCAACAGACGAGGTCCCCAGTACC 2


RESULT   8
US-09-978-917A-25/c
; Sequence 25, Application US/09978917A
; GENERAL INFORMATION:
;  APPLICANT: Maxygen Aps; Maxygen Holdings
;  TITLE OF INVENTION: Protein C or activated protein C-like molecules
;  FILE REFERENCE: 0219us310 - protein C
;  CURRENT APPLICATION NUMBER: US/09/978,917A
;  CURRENT FILING DATE:  2001-10-17
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-978-917A-25


  Query Match             1.1%;  Score 20.6;  DB 5;  Length 49;
  Best Local Similarity   67.4%;  Pred. No. 5.2e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    713 gccccagcgggcaccctggacttggtgcccggcgggctgacct 755
          |||||   || || | ||| |||| || || || ||   ||||
Db     46 GCCCCGCAGGCCAGCTTGGTCTTGTTGTCCAGCAGGACCACCT 4


RESULT   9
US-09-997-623-25/c
; Sequence 25, Application US/09997623
; GENERAL INFORMATION:
;  APPLICANT: Maxygen Aps; Maxygen Holdings
;  TITLE OF INVENTION: Protein C or activated protein C-like molecules
;  FILE REFERENCE: 0219us410 - protein C
;  CURRENT APPLICATION NUMBER: US/09/997,623
;  CURRENT FILING DATE:  2001-11-29
;  PRIOR APPLICATION NUMBER: US 09/978,917
;  PRIOR FILING DATE: 2001-10-17
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-997-623-25


  Query Match             1.1%;  Score 20.6;  DB 5;  Length 49;
  Best Local Similarity   67.4%;  Pred. No. 5.2e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    713 gccccagcgggcaccctggacttggtgcccggcgggctgacct 755
          |||||   || || | ||| |||| || || || ||   ||||
Db     46 GCCCCGCAGGCCAGCTTGGTCTTGTTGTCCAGCAGGACCACCT 4


RESULT  10
US-09-745-237A-3833
; Sequence 3833, Application US/09745237A
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
;  FILE REFERENCE: 400/007 (MBHB00-918-A)
;  CURRENT APPLICATION NUMBER: US/09/745,237A
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 4550
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3833
;   LENGTH: 48
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-745-237A-3833


  Query Match             1.1%;  Score 20.4;  DB 5;  Length 48;
  Best Local Similarity   60.9%;  Pred. No. 5.7e+05;
  Matches   28;  Conservative    2;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    137 ggctgggggtaggggactccctacaggcacgcagccctgagaccgc 182
          ||| ||||| |||  ||:|||: || | ||   | || |  |  ||
Db      1 ggcaggggggaggaaacucccuucaaggacaucguccgggaagggc 46


RESULT  11
US-60-384-980-69/c
; Sequence 69, Application US/60384980
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: D0211 PSP
;  CURRENT APPLICATION NUMBER: US/60/384,980
;  CURRENT FILING DATE:  2002-06-03
;  NUMBER OF SEQ ID NOS: 1219
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
```

```
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-384-980-69


  Query Match             1.1%;  Score 20.2;  DB 8;  Length 41;
  Best Local Similarity   68.3%;  Pred. No. 6.2e+05;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1233 acaaggagaagggcatcttcaaaattgaggactcagcccag 1273
          | | || |||| | || ||  ||  ||||| ||||| |||
Db     41 AAAGGGGGAAGAGAATGTTACAATCTGAGGTCTCAGGGCAG 1


RESULT  12
US-10-099-926-1257/c
; Sequence 1257, Application US/10099926
; GENERAL INFORMATION:
;  APPLICANT: King, Gordon E.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Secrist, Heather
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.547C2
;  CURRENT APPLICATION NUMBER: US/10/099,926
;  CURRENT FILING DATE:  2002-03-17
;  NUMBER OF SEQ ID NOS: 1982
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1257
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-099-926-1257


  Query Match             1.1%;  Score 20.2;  DB 7;  Length 50;
  Best Local Similarity   63.3%;  Pred. No. 6.2e+05;
  Matches   31;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy    473 cccgacacggtgtcgcggacaggcttggagaaggcggcagcgggggcag 521
          | |||||| |   || |||| || | || |  || | | |  ||| |||
Db     49 CACGACACAGCAACGTGGACGGGATGGGGGCGGGGGCCTGGAGGGCCAG 1


RESULT  13
US-09-745-237A-3872
; Sequence 3872, Application US/09745237A
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
;  FILE REFERENCE: 400/007 (MBHB00-918-A)
;  CURRENT APPLICATION NUMBER: US/09/745,237A
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 4550
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3872
;   LENGTH: 48
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-745-237A-3872


  Query Match             1.1%;  Score 20;  DB 5;  Length 48;
  Best Local Similarity   61.4%;  Pred. No. 6.8e+05;
  Matches   27;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    138 gctgggggtaggggactccctacaggcacgcagccctgagaccg 181
          || ||||| ||| ||:|||: || | ||   | || |   |||
Db      2 gcaggggggaggaaacucccuucaaggacaucguccgggagccg 45


RESULT  14
US-09-745-237A-3877
; Sequence 3877, Application US/09745237A
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  TITLE OF INVENTION: Method and Reagent for the Treatment of Alzheimer's Disease
;  FILE REFERENCE: 400/007 (MBHB00-918-A)
;  CURRENT APPLICATION NUMBER: US/09/745,237A
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 4550
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3877
;   LENGTH: 48
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-745-237A-3877


  Query Match             1.1%;  Score 20;  DB 5;  Length 48;
  Best Local Similarity   61.4%;  Pred. No. 6.8e+05;
  Matches   27;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    139 ctgggggtaggggactccctacaggcacgcagccctgagaccgc 182
          | ||||| |||  ||:|||: || | ||   | || |   ||||
Db      3 ccgggggggaggaaacucccuucaaggacaucguccgggagccgc 46


RESULT  15
US-09-913-238-45/c
; Sequence 45, Application US/09913238
; GENERAL INFORMATION:
;  APPLICANT: Barbas, Carlos F.
;  APPLICANT:  Steinberger, Peter
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR
;  TITLE OF INVENTION:  INHIBITING CCR5-DEPENDENT INFECTION OF CELLS BY HIV-1
;  FILE REFERENCE: TSRI 728.1
;  CURRENT APPLICATION NUMBER: US/09/913,238
;  CURRENT FILING DATE:  2002-03-05
;  PRIOR APPLICATION NUMBER: PCT/EPO 0/12419
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: US 60/169,653
;  PRIOR FILING DATE: 1999-12-08
;  NUMBER OF SEQ ID NOS: 85
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Synthesized
US-09-913-238-45


  Query Match             1.0%;  Score 19.8;  DB 5;  Length 44;
  Best Local Similarity   69.2%;  Pred. No. 7.4e+05;
  Matches   27;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    358 cagctccctgggcaccaggcagctaacagacacagccgc 396
          | ||  ||||||| ||  ||||||| ||  | |  ||||
Db     39 CGGCCGCCTGGGCCCCTCGCAGCTAGCAAGCTCCTCCGC 1


Search completed: August  6, 2002, 23:42:38
```

Job time: 7651 sec